A multiscale study of stochastic spatially-extended conductance-based models for excitable systems

Alexandre Genadot

École Doctorale de Sciences Mathématiques de Paris Centre

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

Discipline : Mathématiques

présentée par

Alexandre GENADOT

pour obtenir le grade de :

DOCTEUR DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

sujet de la thèse :

Étude multi-échelle de modèles probabilistes pour les systèmes excitables avec composante spatiale.

soutenue le 4 novembre 2013

devant la commission d'examen composée de:

M.	Arnaud DEBUSSCHE	ENS Cachan	Rapporteur
M.	Amaury LAMBERT	Université Paris 6	Examinateur
M.	Gabriel LORD	Heriot-Watt University	Rapporteur
M.	Florent MALRIEU	Université de Tours	Examinateur
Mme	Michèle THIEULLEN	Université Paris 6	Directrice de thèse

Laboratoire de Probabilités et Modèles Aléatoires
Université Pierre et Marie Curie (Paris 6)
Tours 16-26, 1er étage
4 place Jussieu
75 005 PARIS

École Doctorale de Sciences Mathématiques de Paris Centre
Université Pierre et Marie Curie (Paris 6)
Tours 15-25, 1er étage
4 place Jussieu
75 005 PARIS

Baissatz-vos montanhas,
Planas, auçatz-vos,
Per que pòsca veire
Mas amors ont son.

Se Canta.

À mon père

Remerciements

Je souhaite adresser mes sincères remerciements à Arnaud Debussche et Gabriel Lord pour avoir rapporté ma thèse et pour l'intérêt qu'ils ont porté à mon travail. Je tiens également à exprimer ma reconnaissance à ma directrice de thèse, Michèle Thieullen, pour sa disponibilité et ses encouragements continus ainsi qu'à Muriel Boulakia dont la collaboration m'a permis d'écrire le sixième et dernier chapitre de cette thèse. Je remercie aussi Amaury Lambert et Florent Malrieu d'avoir accepté de participer à mon jury.

Ces trois années de doctorat auraient été bien différentes sans le piquant des membres du 203 : Alexandre B., Antoine, Benjamina, Bruno, Clara, Florian, Franck, Weibing et Xan. Je ne saurais citer sans en oublier tous les membres du LPMA qui ont rendu ces trois années des plus agréables. Permettez-moi simplement d'exprimer ma gratitude à l'égard de nos charmantes secrétaires, notre estimable bibliothécaire, notre irremplaçable informaticien ainsi que les thésards passés et présents des 110, 131, 201 et 225.

De Montpellier à Paris, en passant par Bordeaux et Toulouse, vous savez bien, mes très chers amis, que votre aide, votre écoute et vos conseils me furent précieux. Je tiens aussi à remercier les membres de ma famille et plus particulièrement ma mère. Merci Maman. Pour tout ce que tu as fait, ce que tu fais et ce que – comment en douter ? – continueras de faire pour moi.

Enfin, merci à toi Lucile, *chatona tant polida...* Tu as fait preuve, durant ces trois ans, d'une patience admirable. Tu m'as soutenu lorsque je n'y croyais plus et inspiré chaque jour une volonté nouvelle.

Résumé

L'objet de cette thèse est l'étude mathématique de modèles probabilistes pour la génération et la propagation d'un potentiel d'action dans les neurones et plus généralement de modèles aléatoires pour les systèmes excitables. En effet, nous souhaitons étudier l'influence du bruit sur certains systèmes excitables multi-échelles possédant une composante spatiale, que ce soit le bruit contenu intrinsèquement dans le système ou le bruit provenant du milieu. Ci-dessous, nous décrivons d'abord le contenu mathématique de la thèse. Nous abordons ensuite la situation physiologique décrite par les modèles que nous considérons.

Pour étudier le bruit intrinsèque, nous considérons des processus de Markov déterministes par morceaux à valeurs dans des espaces de Hilbert (*Hilbert-valued PDMP*). Nous nous sommes intéressés à l'aspect multi-échelles de ces processus et à leur comportement en temps long.

Dans un premier temps, nous étudions le cas où la composante rapide est une composante discrète du PDMP. Nous démontrons un théorème limite lorsque la composante rapide est infiniment accélérée. Ainsi, nous obtenons la convergence d'une classe de *Hilbert-valued PDMP* contenant plusieurs échelles de temps vers des modèles dits moyennés qui sont, dans certains cas, aussi des PDMP. Nous étudions ensuite les fluctuations du modèle multi-échelles autour du modèle moyenné en montrant que celles-ci sont gaussiennes à travers la preuve d'un théorème de type *central limit*.

Dans un deuxième temps, nous abordons le cas où la composante rapide est elle-même un PDMP. Cela requiert de connaître la mesure invariante d'un PDMP à valeurs dans un espace de Hilbert. Nous montrons, sous certaines conditions, qu'il existe une unique mesure invariante et la convergence exponentielle du processus vers cette mesure. Pour des PDMP dits diagonaux, la mesure invariante est explicitée. Ces résultats nous permettent d'obtenir un théorème de moyennisation pour des PDMP « rapides » couplés à des chaînes de Markov à temps continu « lentes ».

Pour étudier le bruit externe, nous considérons des systèmes d'équations aux dérivées partielles stochastiques (EDPS) conduites par des bruits colorés. Sur des domaines bornés de $\mathbb{R}^2$ ou $\mathbb{R}^3$, nous menons l'analyse numérique d'un schéma de type différences finies en temps et éléments finis en espace. Pour une classe d'EDPS linéaires, nous étudions l'erreur de convergence forte de notre schéma. Nous prouvons que l'ordre de convergence forte est deux fois moindre que l'ordre de convergence faible. Par des simulations, nous montrons l'émergence de phénomènes d'ondes ré-entrantes dues à la présence du bruit dans des domaines de dimension deux pour les modèles de Barkley et Mitchell-Schaeffer.

L'étude mathématique décrite précédemment s'inspire de modèles déterministes classiques pour les milieux excitables, principalement les neurones et les cellules cardiaques. Il s'agit de modèles « à conductances ». Les plus utilisés dans cette thèse sont les modèles de Hodgkin et Huxley, de Barkley et de Mitchell-Schaeffer. Ces modèles déterministes décrivent l'évolution combinée en temps et en espace du potentiel trans-membranaire d'une cellule isolée et l'état des canaux ioniques situés sur la membrane. Ce phénomène électro-chimique met en jeu des échelles de temps différentes. Pour cette raison les modèles sont multi-échelles. De plus, ils correspondent à une évolution limite lorsque le nombre de canaux est très grand. Les modèles probabilistes de type PDMP utilisés dans la thèse sont plus réalistes sur le plan biologique puisqu'ils correspondent à une cellule ayant un nombre fini de canaux. Le modèle de type EDPS correspond à une cellule non isolée soumise au bruit ambiant.

Les théorèmes obtenus sur les *Hilbert-valued PDMP* avec plusieurs échelles de temps permettent de diminuer la dimension de ces modèles ou de les remplacer par des approximations diffusions. Cela permet d'aborder plus facilement la simulation ainsi que l'analyse mathématique et le couplage de ces modèles. L'étude menée sur les systèmes de type EDPS est motivée par la recherche de phénomènes d'arythmie. En effet, ces phénomènes sont encore mal expliqués mathématiquement.

Mots clés : Processus de Markov déterministes par morceaux ; Équations aux dérivées partielles stochastiques ; Systèmes multi-échelles ; Modèles de neurones ; Modèles de cellules cardiaques ; Moyennisation ; Schémas numériques.

Abstract

The purpose of the present thesis is the mathematical study of probabilistic models for the generation and propagation of an action potential in neurons and more generally of stochastic models for excitable cells. Indeed, we want to study the effect of noise on multiscale spatially extended excitable systems. We address the intrinsic as well as the extrinsic source of noise in such systems. Below, we first describe the mathematical content of the thesis. We then consider the physiological situation described by the considered models.

To study the intrinsic or internal noise, we consider Hilbert-valued Piecewise Deterministic Markov Processes (PDMPs). We are interested in the multiscale and long time behavior of these processes.

In a first part, we study the case where the fast component is a discrete component of the PDMP. We prove a limit theorem when the speed of the fast component is accelerated. In this way, we obtain the convergence of a class of Hilbert-valued PDMPs with multiple timescales toward so-called averaged processes which are, in some cases, still PDMPs. Then, we study the fluctuations of the multiscale model around the averaged one and show that the fluctuations are Gaussians through the proof of a Central Limit Theorem.

In a second part, we consider the case where the fast component is itself a PDMP. This requires knowledge about the invariant measure of Hilbert-valued PDMPs. We show, under some conditions, the existence and uniqueness of an invariant measure and the exponential convergence of the process toward this measure. For a particular class of PDMPs that we call diagonals, the invariant measure is made explicit. This, in turn, allow us to obtain averaging results for "fast" PDMPs fully coupled to "slow" continuous time Markov chains.

To study the extrinsic or external noise, we consider systems of Stochastic Partial Differential Equations (SPDEs) driven by colored noises. On bounded domains of $\mathbb{R}^2$ or $\mathbb{R}^3$, we analyze numerical schemes based on finite differences in time and finite elements in space. For a class of linear SPDEs, we obtain the strong error of convergence of such schemes. For simulations, we show the emergence of re-entrant patterns due to the presence of noise in spatial domains of dimension two for the Barkley and Mitchell-Schaeffer models.

The mathematical study described above is inspired by classical deterministic models for excitable media, especially for neural and cardiac cells. They are conductance based models. In the present work, the underlying deterministic models are mostly the Hodgkin-Huxley, Barkley and Mitchell-Schaeffer models. These models describe the evolution in time and space of the trans-membrane potential of an isolated cell as well as the evolution of the states of the ionic channels

in the membrane. This electro-chemical phenomenon brings into play different timescales. This is why the considered models have multiple timescales. Moreover, they correspond to an evolution at the limit for a large number of ionic channels. The PDMP models used in the present thesis are more biologically relevant since they correspond to cells with a finite number of ionic channels. The SPDE models correspond to a cell submitted to external noise.

The theorems obtained for Hilbert-valued PDMPs with multiple timescales show how to decrease the dimension of the models or how to approximate them by diffusions. This allows more tractable simulations, mathematical analysis and coupling of such models. The study of systems of SPDEs is motivated by the search for arrhythmia in cardiac cells. Indeed, this phenomenon is still hard to explain mathematically.

Keywords: Piecewise deterministic Markov processes; Stochastic partial differential equations; Neuron models; Cardiac models; Multiscale systems; Averaging; Numerical schemes.

Contents

Chapter 1

Introduction

This document gathers the works I achieved as a PhD student under the supervision of Michèle Thieullen at the Laboratoire de Probabilités et Modèles Aléatoires of the Université Pierre et Marie Curie (Paris 6).

The central subject of the present work is the study of stochastic models for excitable media. This thesis is composed of six chapters. After the introduction (Chapter 1) and a chapter collecting some technical results (Chapter 2), Chapters 3 and 4 deal with the reduction of Hilbert-valued Piecewise Deterministic Markov Processes (PDMPs) for excitable cells with multiple timescales. Then, we study the long time behavior of Hilbert-valued PDMPs with application to averaging (Chapter 5). At last, Chapter 6 investigates the effect of noise on excitable media through numerical methods. These chapters are essentially made up of a published paper [GT12] and four submitted papers [ABG+13, BGT13, GT13a, GT13b].

The introduction is organized as follows. Section 1.1 is about the mathematical modeling of excitable cells. We start with the basic properties of excitable systems (Section 1.1.1) and then present the basic physiology of neurons (Section 1.1.2). We go on with the description of conductance-based models (Section 1.1.3) and phenomenological models (Section 1.1.4). In Section 1.2, we present the mathematical representations of these two kinds of models. Hilbert-valued piecewise deterministic Markov models (Section 1.2.1) and stochastic partial differential equations (Section 1.2.2) are considered. Section 1.3 describes, chapter by chapter, the main mathematical results obtained in this thesis. Some perspectives related to our work are discussed in Section 1.4.

1.1 Mathematical modeling of excitable cells

1.1.1 Common facts for excitable systems

In this paragraph, we present the main features of excitable systems. Such systems have been reported in different areas of physics, chemistry and life sciences. The neural cell, one of the most important example of biological excitable system, is described in the next section.

We proceed to the formal description of an excitable system. A dynamical system (S) is said to be excitable when

- the system (S) possesses a single stable *rest state*. More generally, this rest state can be simply an equilibrium point or a small-amplitude limit cycle,
- the system (S) can leave the resting state, which is dynamically stable, only if a sufficiently strong external perturbation is applied. In this case, the large excursions from the resting state are often called *spikes*, and their occurrence is referred to as a firing. There is thus a *threshold* of activation from which the system begins to spike. When the system's variable values are superthreshold, the system is said to be *excited*. Otherwise, it is said to be *quiescent* or *unexcited*,
- the system is *refractory* after a spike, which means that it takes a certain recovery time before another excitation can evoke a second spike.

These properties are illustrated in Figure 1.1. Besides having an interesting temporal behavior, when a large number of excitable systems are coupled to one another, they exhibit a rich variety of spatio-temporal behaviors, depending on the strength and topology of the coupling between them. Phenomena such as pulse and spiral wave propagation, scroll waves, localized spots, periodic patterns in space and/or in time, and spatio-temporal chaos have been recorded in a wide range of physical systems. In perspective, one of the most remarkable properties of excitable systems is their ability to synchronize, both among themselves or to external periodic (or more complex) signals.

For more details, we refer the reader to the review [LGONSG04] on the effect of noise in excitable media.

1.1.2 An example of excitable cells: neurons

The aim of this paragraph is not to give an exhaustive and detailed presentation of neurons, but rather to familiarize the reader with basic notions concerning their bio-physiology in order to have a better understanding of the models presented

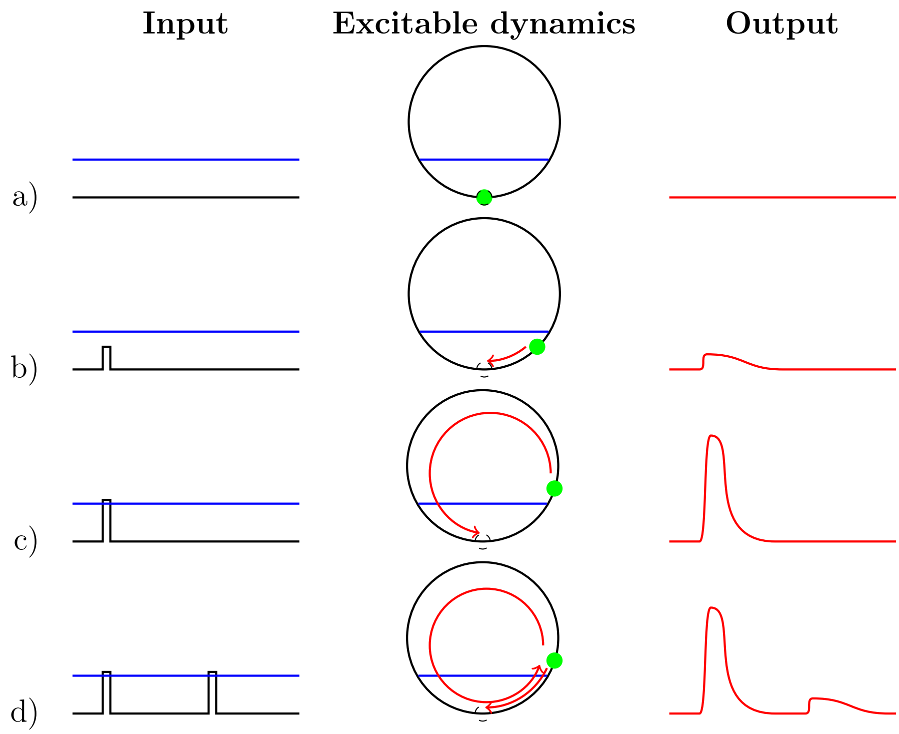


Figure 1.1: Features of excitable systems. This figure follows Figure 1 of [LGONSG04]. Different kinds of inputs (left column) cause different kinds of responses (right column) of the excitable dynamics (middle column). a) No input: system at rest. b) An input below threshold (blue line) results in a small amplitude motion around the system stable state. c) An input exceeding the threshold leads to a large-amplitude excursion of the system variables (spike). d) If the pulses are too close, the system does not respond noticeably to the second perturbation because of refractory effects.

later in the text. This paragraph is largely inspired by [SBB$^+$12], Chapter 6, for the description of the neuron and the mechanisms of generation and propagation of an action potential by the neural cell. Another classical reference used in the text is [Hil84] about ion channels of excitable membranes.

Neurons and *glia* are the cellular building blocks of the nervous system. Neurons are interconnected, highly differentiated bio-electrically driven cellular units of the nervous system supported by the more numerous cells termed glia. Morphologically, in a typical neuron, three major regions can be defined:

- the *cell body*, also referred to as the *soma*, which contains the nucleus and the major cytoplasmic organelles (the 'organs' of the cell body),

- a variable number of *dendrites*, which emanate from the soma and differ in size and shape, depending on the neuronal type. Dendrites exhibit a tree structure.
- a single *axon*, which, in most cases, extends much farther from the cell body than the dendritic tree.

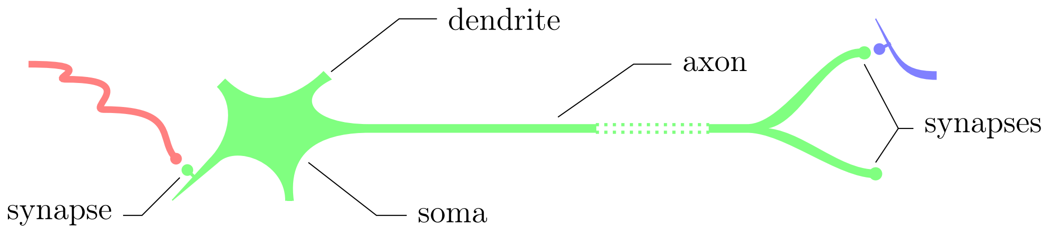


Figure 1.2: Schematic representation of a neuron

The cell body and dendrites are the two major domains where the neural cell receives inputs from other neurons or muscles or peripheral organs. The axon, at the other pole of the neuron, is responsible for transmitting neural information, that is the inputs, to interconnected target neurons. This electrical signal is called the nerve impulse or the *action potential* or the *spike*. At the interfaces of axon terminals with target cells are the *synapses* which represent specialized zones of contact with a dendrite or the soma of other neurons.

Like in any biological cell, the neuronal membrane exhibits a difference of potential between the internal and the external cellular media which is called trans-membrane potential, membrane potentials or simply *potential*. This membrane potential is generated by the unequal distributions of ions, particularly potassium (K^+), sodium (Na^+) and chloride (Cl^-), across the plasma membrane. The resting membrane is permeable to these ions and their inhomogeneous distribution between the inside and the outside of the cell is maintained by ionic pumps and exchangers.

In most neurons, action potentials are initiated in the initial portion of the axon, known as the axon initial segment. Once a spike is initiated, it propagates along the axon in its normal direction to the synaptic terminals. Axons have a threshold for the initialization of an action potential. That is, the neuron generates a spike if the input is above a certain value termed the threshold. It is this property that makes a neuron an excitable cell, see Section 1.1.1, a small variation in the input may result in a dramatic variation of the membrane potential. Two major phenomena are responsible for the generation of an action potential:

- an increase in the voltage of the axon membrane produces a large but transient flow of positive charges carried by Na^+ ions flowing into the cell (inward current),

- this transient inward current is followed by a sustained flow of positive charges out of the cell (the outward current) carried by a sustained flux of K^+ ions moving out of the cell.

These flows of ions across the plasma membrane take place in ionic channels. These channels are generally permeable to one specific ion's specie and are present in variable density at different locations in the membrane. A channel can be, roughly speaking, open or closed. When a spike is initiated at the axon initial segment, the depolarization of the membrane increases the probability of Na^+ channels being in the open state. This causes the flow of Na^+ entering the cell, depolarizing the axon and opening still more Na^+ channels, causing yet more depolarization of the membrane until the resting potential of the Na^+ ions is reached. In a slightly delayed time, the depolarization of the membrane potential increases the probability of K^+ channels being open and allows positive charges to exit the cell. Then, at some point, the outward movement of K^+ ions dominates and the membrane potential is re-polarized, corresponding to the decrease of the action potential. The persistence of the outward current for a few milliseconds following the spike generates the after-hyper-polarization. During this phase of hyper-polarization, the ionic channels get back to their resting states, preparing the axon for generation of the next action potential. We stress out that the opening and closing of the ionic channels (often referred to as conformal changes) during the generation of a spike is a *voltage-dependent* phenomenon: the conformal changes of a ionic channel depends on the membrane potential.

1.1.3 Conductance-based models

The understanding of bio-electrical properties of neurons arose in the nineteenth century from the merging of two different domains: on the one hand the study of current spread in nerve cells and muscles, on the other hand the development of cable theory for long distance transmission of electric current through cables on the ocean floor. The communication of information between neurons and between neurons and muscles or peripheral organs requires that signals travel over considerable distances (with respect to the size of the dendrites or the cell body). The cable theory was first applied to the nervous system in the late nineteenth century to model the flow of electric current through nerve fibers. By the 1930s and 1940s, it was applied to the axon of simple invertebrates (crab and squid). This was the first step toward the development of the Hodgkin-Huxley equations [HH52] for the action potential in an axon (1952). Hodgkin and Huxley received the 1963 Nobel prize of physiology and medicine for their work. We present in this section the deterministic Hodgkin-Huxley equations and then a mathematical generalization based on the works [Aus08, FWL05]. This generalization takes into account the intrinsic stochasticity of ion channels. We choose the common denomination of conductance-based neuron model for what is sometimes also called generalized Hodgkin-Huxley model in the literature. Conductance-based models without spatial propagation are very popular, see the review [Rin90], section 4. They are called point models. Conductance-based models are also widely used to model cardiac cells, see [DIP+10, Sac04]. Indeed, similar mechanisms than those described in Section 1.1.2 for neurons lead to the generation and propagation of cardiac potential. Existence and uniqueness of a solution to the Hodgkin-Huxley equations including spatial propagation have been obtained in [ES70, Lam86, RK73]. Only a few mathematical papers are available about the spatially extended stochastic conductance-based models we are going to present in this section. The first one is certainly due to [Aus08] although simulation studies have been performed earlier [FWL05].

As explained in Section 1.1.2, in most neural cells, the action potential is initiated in the axon initial segment. The axon is often considered as a cable much longer than larger. It is therefore modeled as a segment I in the sequel.

The deterministic Hodgkin-Huxley model

We proceed to a brief description of the deterministic model introduced by Hodgkin and Huxley in [HH52]. The Hodgkin-Huxley model is a system of four coupled partial differential equations describing the evolution of the membrane potential at a given point of the axon and the probability of ionic channels to be open. This system reads as follows

$$\left\{\begin{array}{rcl} C\partial_t u & = & \frac{a}{2R}\partial_{xx}u + c_{\mathrm{Na}}m^3h(v_{\mathrm{Na}} - u) + c_{\mathrm{K}}n^4(v_{\mathrm{K}} - u) + c_{\mathrm{l}}(v_{\mathrm{l}} - u), \\ \partial_t y & = & (1-y)a_y(u) - yb_y(u), \quad y = m, n, h. \end{array}\right. \tag{1.1}$$

Boundary conditions associated to these equations will be given later in the text, when needed. The variable $u_t(x)$ denotes the membrane potential at position $x \in I$ on the axon at time t, that is the difference of potential between the inside and the outside of the neuron. Only sodium, potassium and leak currents are considered in this model. An ionic channel is composed of four gates: three gates of type m and one of type h for a sodium channel and four gates of type n for a potassium channel. In the model (1.1), these gates may be open or closed. A specific gate opens and closes at voltage dependent rates a_y and b_y, for $y = m, n, h$ respectively. When open, that is when all its gates are open, a channel allows a current to pass with conductance c_i with $i =$ Na or K. C is the capacity of the axon membrane, R its internal resistance and a the radius of the axon. The quantities m^3h and n^4 may be interpreted as the probability of open sodium and potassium channels respectively or in other word as the inward and the outward currents respectively. Figure 1.3 displays, at a given point $x \in I$ of the axon, the evolution of the action potential u and the inward and outward currents m^3h and n^4 with respect to time according to the Hodgkin-Huxley model (1.1). The coupling of the inward and outward currents to the evolution of the action potential described above is clearly visible in this figure. As an illustration of propagation of an action potential, Figure 1.4 displays simulations of the propagation of the action potential in space and time according to the Hodgkin-Huxley system when the axon is assimilated to a segment.

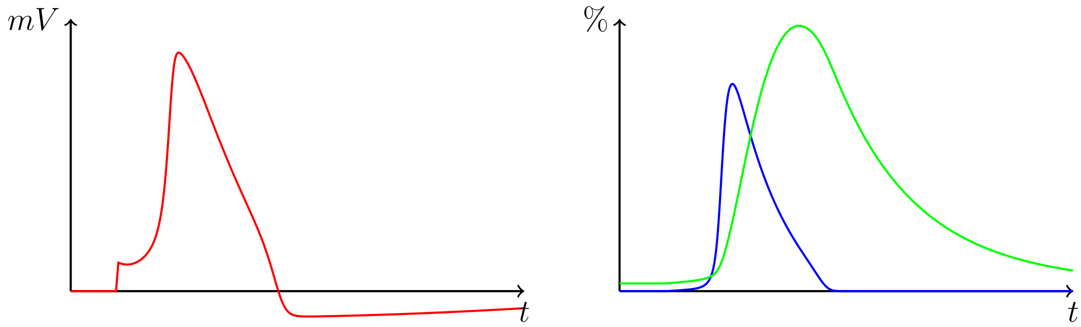


Figure 1.3: Action potential (red) and inward (blue) and outward (green) currents in the deterministic Hodgkin-Huxley model described by equations (1.1). The scale for the x-axis is the same for both graphs: the system (1.1) is displayed between the times 0 and 10 in ms. For the y-axis, the inward and outward currents are in $[0, 1]$ like probabilities whereas the action potential evolves between -20 and 120 mV.

Stochastic conductance-based models with a finite number of ion channels

We go on with the presentation of a class of stochastic models from which generalized Hodgkin-Huxley models will emerge . This class describes the evolution of an action potential at the scale of ion channels. The influence of noise from voltage-gated ion channels on the generation and propagation of an action potential has been highlighted in [CW96, FL07, FWL05, FL94, WKAK98, WRK00]. Stochastic conductance based models are studied in Chapters 3 and 4 of the present work from the perspective of stochastic averaging.

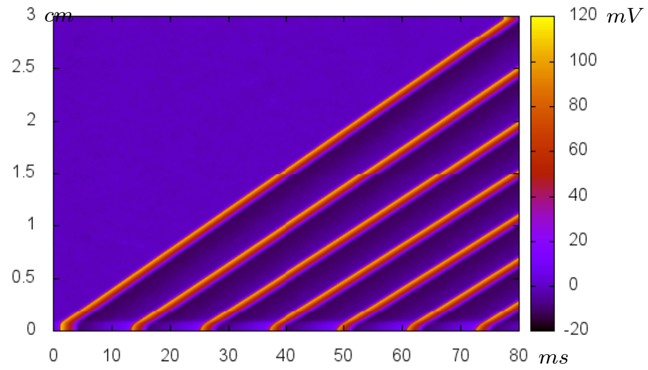


Figure 1.4: Propagation of an action potential according to the deterministic Hodgkin-Huxley model (1.1). Space and time are respectively given by the y and x-axis. The Hodgkin-Huxley system is excited on the initial segment on the axon by a constant superthreshold stimulation. As a consequence, a train of spikes propagating along the nerve fiber is observed.

All along the axon cable are the ionic channels in various densities at different locations of the axon. We assume that the ionic channels are in finite numbers and located at discrete sites z_i through the axon membrane for $i \in \mathcal{N}$ where $\mathcal{N}$ is a finite set. A specific channel can be in several states. Actually, a channel can be in an activated, inactivated or deactivated state where for example, the inactivated state can in turn correspond to different states. Moreover, an ionic channel is permeable to only one specific ion specie, but mathematically, taking into account this specificity only changes the notations and we consider in the general model that a channel is permeable to all kinds of ion species. What is mathematically relevant is that a specific ionic channel can be in a finite number of states. We denote this state space by E and an element of this space by the Greek letters ξ or ζ. When an ionic channel is open, it allows a flow of ions to enter or leave the cell, that is, it allows a current to pass. When open, an ionic channel in the state ξ allows the ions of a certain specie associated to ξ to pass through the membrane with a driven potential v_ξ. This driven potential governs the direction of the ionic flux: it says whether the current is inward or outward. A channel possesses also its proper internal resistance according to its state. As for system (1.1), for notation purposes, we prefer to work with the inverse of this resistance, namely the conductance, which is denoted in the sequel by c_ξ for an ionic channel in state ξ.

As above, $u_t(x)$ denotes the membrane potential at position x of the axon at time t. We denote by $r_t(i)$ the state of the ionic channel at position z_i at time t. According to stochastic conductance-based neuron models we are going to work with, the evolution of the membrane potential is governed by the following cable equation

$$C\partial_t u_t = \partial_x(\alpha\partial_x u_t) + \frac{1}{|\mathcal{N}|}\sum_{i\in\mathcal{N}} c_{r_t(i)}(v_{r_t(i)} - \Phi_i(u_t))\Psi_i. \tag{1.2}$$

In full generality C, the internal capacitance of the membrane, may depend on the space variable x. The function α depends on some bio-physical constants of the axon like its diameter and internal resistance which may both depend on the spatial position. The function Φ_i indicates in which manner the local potential of the membrane around the location z_i is affected by the opening of the channel at this location. For instance, when this influence is very localized, $\Phi_i(u_t) = u_t(z_i)$. Then in this case Ψ_i is equal to δ_{z_i}, the Dirac mass at point z_i. One may also consider $\Phi_i(u_t) = (u_t, \phi_{z_i})$ where ϕ_{z_i} are mollifiers. Here $(\cdot,\cdot)$ denotes the usual scalar product in $L^2(I)$. Then, in this case Ψ_i is equal to ϕ_{z_i}. The biological meaning of considering mollifiers instead of Dirac mass is that when the channel located at z_i is open and allows a current to pass, not just the voltage at z_i is affected, but also the voltage on a small area around z_i (see [BR11]). More generally, Ψ_i indicates which portion of the ion channel is affected by the opening of the ion channel at position z_i. Equation (1.2) tells us that the membrane current is the sum of all the ionic currents and that this current is propagated along the axon thanks to the diffusive operator $\partial_x(\alpha\partial_x u_t)$.

As previously mentioned, an ion channel can be in a finite number of different states. Therefore, we have to describe the mechanism of passing from one state to another. If the membrane potential were held fixed, this mechanism would simply follow the kinetic scheme of a classical Continuous Time Markov Chain (CTMC). However, the membrane potential evolves with time and the rate of jumps from one state to another for an ionic channel is voltage-dependent. This leads to a non-homogeneous evolution in time for the state of the ionic channel at position z_i according to the following dynamic:

$$\mathbb{P}(r_{t+h}(i) = \zeta | r_t(i) = \xi) = \alpha_{\xi\zeta}(\Phi_i(u_t))h + \mathrm{o}(h). \tag{1.3}$$

That is, the jump process $(r_t(i), t \geq 0)$ goes from state ξ to state ζ at rate $\alpha_{\xi\zeta}$ where the latter is a voltage-dependent function. Moreover, the evolution of a given ionic channel $r(i)$ is often assumed to be independent from the evolution of the other channels over infinitesimal timescale. More precisely, this means that the coordinate processes $(r_{t+h}(i))_{h>0}$ for $i \in \mathcal{N}$ are independent conditionally on

$\mathcal{F}_t$ to first order in h as h goes to zero (we denote by $\mathcal{F}_t$ the σ-algebra generated by the process (u, r) up to time t).

$$\begin{array}{ccccccc}
m_0h_0 & \underset{b_m}{\overset{3a_m}{\rightleftharpoons}} & m_1h_0 & \underset{2b_m}{\overset{2a_m}{\rightleftharpoons}} & m_2h_0 & \underset{3b_m}{\overset{a_m}{\rightleftharpoons}} & m_3h_0 \\
b_h \upharpoonleft\downharpoonright a_h & & b_h \upharpoonleft\downharpoonright a_h & & b_h \upharpoonleft\downharpoonright a_h & & b_h \upharpoonleft\downharpoonright a_h \\
m_0h_1 & \underset{b_m}{\overset{3a_m}{\rightleftharpoons}} & m_1h_1 & \underset{2b_m}{\overset{2a_m}{\rightleftharpoons}} & m_2h_1 & \underset{3b_m}{\overset{a_m}{\rightleftharpoons}} & \mathbf{m_3h_1}
\end{array}$$

Kinetic of a sodium ionic channel

$$n_0 \underset{b_n}{\overset{4a_n}{\rightleftharpoons}} n_1 \underset{2b_n}{\overset{3a_n}{\rightleftharpoons}} n_2 \underset{3b_n}{\overset{2a_n}{\rightleftharpoons}} n_3 \underset{4b_n}{\overset{a_n}{\rightleftharpoons}} \mathbf{n_4}$$

Kinetic of a potassium ionic channel

Figure 1.5: Kinetic schemes for Sodium and Potassium ionic channels in a generalized Hodgkin-Huxley model. In this model, a sodium channel may be in 8 different states in the set $E_{\mathrm{Na}} = \{m_0h_0, m_1h_0, m_2h_0, m_3h_0, m_0h_1, m_1h_1, m_2h_1, m_3h_1\}$ where m_3h_1 codes for the open state. In the same way, a potassium channel may be in 5 different states in the set $E_{\mathrm{K}} = \{n_0, n_1, n_2, n_3, n_4\}$ where n_4 codes for the open state.

As an example, in [Aus08], the author assumes that the axon can be assimilated to the segment $I = [-l, l]$ and that the ionic channels are at locations $z_i = \frac{i}{N}$ for $i \in \mathcal{N} = \mathbb{Z} \cap N(-l, l)$. This same distribution is used in the present thesis in Chapters 3 (all along the chapter) and 4 (as an example). The equation describing the evolution of the membrane potential for the corresponding model becomes

$$\partial_t u_t = \partial_{xx} u_t + \frac{1}{N} \sum_{i \in \mathcal{N}} c_{r_t(i)}(v_{r_t(i)} - u_t(z_i))\delta_{z_i}. \tag{1.4}$$

with relevant initial and boundary conditions described later in the text. The dynamic of the ionic channels is given by

$$\mathbb{P}(r_{t+h}(i) = \zeta | r_t(i) = \xi) = \alpha_{\xi\zeta}(u_t(z_i))h + \mathrm{o}(h) \tag{1.5}$$

for each channel at location $i \in \mathcal{N}$. In this case, the following theorem has been derived under natural hypotheses.

Theorem 1.1.1 ([Aus08]). *When the number of channels increases to infinity, the stochastic conductance-based model given by equations (1.4-1.5) converges in probability, in an appropriate space, towards a deterministic conductance based model where the evolution equation for the membrane potential u_t is*

$$\partial_t u_t = \partial_{xx} u_t + \sum_{\xi \in E} c_\xi (v_\xi - u_t) p_{\xi,t}, \tag{1.6}$$

and the evolution of the conductances $p_{\xi,t}$ is given by

$$\frac{\mathrm{d}p_{\xi,t}}{\mathrm{d}t} = \sum_{\zeta \neq \xi} \alpha_{\xi\zeta}(u_t) p_{\zeta,t} - \alpha_{\zeta\xi}(u_t) p_{\xi,t}, \quad \xi \in E. \tag{1.7}$$

Equations (1.6) and (1.7) define a spatially extended deterministic conductance-based model often called the spatially extended generalized Hodgkin-Huxley model. The evolution equation (1.6) on u describes the evolution of the membrane potential along the axon segment over time. For $\xi \in E$, equation (1.7) gives the time evolution of the probability p_ξ that the ionic channels are in state ξ, when the number of ionic channels is infinite. As far as we know, Theorem 1.1.1 was the first mathematical result on stochastic spatially extended conductance-based neuron models obtained in the literature. As pointed out in the title of [Aus08], this result shows the emergence of the deterministic Hodgkin-Huxley model as a limit from the underlying ionic channel mechanism. This result is of first importance because it confirms the consistency of the deterministic and stochastic descriptions of the nerve equations. Generalization of Theorem 1.1.1 has been obtained in [RTW12]: the authors derive a law of large numbers for a general class of models, called compartmental models, which establishes a connection to deterministic macroscopic models as in Theorem 1.1.1. Moreover, they obtained an associated martingale central limit theorem which connects the stochastic intrinsic fluctuations around the deterministic limiting process to diffusion processes.

1.1.4 Phenomenological models

Conductance-based neuron models are biologically relevant since they are based on the actual bio-physical properties of neurons. They describe, at the scale of the ionic channels, the mechanisms of generation and propagation of an action potential. However, these detailed models may present some issues for further analysis. The models presented in the previous section may be of high dimension: the classical deterministic Hodgkin-Huxley model is a set of four partial differential equations (PDEs). High dimensional systems of PDEs (or even ODEs) are difficult to handle theoretically and reductions or simplifications of the original model are precious. The same issues arise when studying networks of neurons: the mathematical model of one single neuron must be as tractable as possible numerically if one wants to simulate a network of thousands of neurons in interaction. To overcome these difficulties, reductions and simplifications of conductance-based models have been proposed. The reduction techniques are much often based on the separation of timescales: among the dynamics leading to the generation of a nerve impulse, some are much faster than others. This is the subject of Chapters 3 and 4 of the present work where reduction techniques are applied to the stochastic conductance-based models presented in Section 1.1.3.

These simplified models, called phenomenological, became very popular to model neurons and cardiac cells. One of the most famous is the FitzHugh-Nagumo system (1962-1969) [Fit55, Fit61, Fit69, NAY62] which can be considered as a paradigm for the description of excitable media. The Fitzhugh-Nagumo model consists in the following system of two evolution equations

$$\left\{\begin{array}{lcl} \partial_t u_t & = & \partial_{xx} u_t + u_t - \frac{1}{3}u_t^3 - w_t, \\ \partial_t w_t & = & \varepsilon(a u_t - b w_t). \end{array}\right. \tag{1.8}$$

Compared to the conductance-based models, the variable $u_t(x)$ is still modeling the membrane potential of a neuron at time t and position x. w is often referred to as the recovery variable of the system, a denomination coming from the dynamical properties that we now proceed to introduce. When an adequate input is added to the first equation in system (1.8), a threshold phenomenon for excitation can be recovered. When excited, the system generates and propagates spikes in a similar way as deterministic conductance-based neuron models. Therefore, the qualitative properties of these more detailed and biologically relevant models are conserved. Moreover, contrary to higher dimensional models, the dynamic of the Fithugh-Nagumo model can be fully described, see for example [Fit69]. One of the main features of this model is that it possesses all the characteristics of excitable systems (c.f. Section 1.1.1), like the conductance-based neuron models presented in the previous section.

Up to now in the present section, we have only considered the Fitzhugh-Nagumo model as a deterministic model. One may want to add noise to this system for several reasons, c.f. [DRL12, GSB11]. Regarding the conductance-based neuron models presented in the previous section, it may be significant to account for the random switching of ion channels by incorporating some noise in the equation for the recovery variable w in system (1.8). One may also consider the random synaptic inputs from other neurons by adding a noise term to the input applied to the variable u of equation (1.8). For the latter, this leads to a stochastic system of the following form:

$$\left\{\begin{array}{lcl}\partial_t u_t & = & \partial_{xx} u_t + u_t - \frac{1}{3}u_t^3 - w_t + \dot{W}_t,\\ \partial_t w_t & = & \varepsilon(a u_t - b w_t),\end{array}\right. \tag{1.9}$$

where $\dot{W}$ is a noise source which will be defined more precisely in Section 1.2.2. In contrast to the deterministic case (1.8), phenomena induced solely by noise may be initiated. Moreover, behaviors displayed in the deterministic setting such as synchronization, resonant behavior and pattern formation are influenced and modified by noise. Noise induced phenomena for finite dimensional slow-fast stochastic differential equations are studied in [BG06]. Let us remark that very recently, the effect of noise on neural networks has received a great deal of attention, see [Bre12, BN13, RB13, FM13, KR13, LS13]. The influence of noise on excitable systems of the form (1.9) is the object of Chapter 6. More precisely, we study the existence of re-entrant patterns in the Barkley [Bar91] and Mitchell-Schaeffer stochastic models [MS03].

1.2 Mathematical tools

The aim of the present section is to introduce the reader to the two main mathematical objects we will work with throughout this thesis: Piecewise Deterministic Markov Processes (PDMPs) and Stochastic Partial Differential Equations (SPDEs).

1.2.1 A class of Piecewise Deterministic Markov Processes

Stochastic conductance-based models of type (1.2-1.3) exhibit a very particular dynamic: the evolution of the membrane potential follows a deterministic PDE whose parameters are randomly updated when a change in the channel states occurs. These models are therefore hybrid models combining a deterministic evolution punctuated by random events. These particular hybrid models are referred to as *switching* PDEs in the sequel and are studied in the framework of Hilbert-valued Piecewise Deterministic Markov Processes (PDMPs). Note that we work with PDMPs without boundaries, that is without deterministic forced jumps.

PDMPs, also called Markovian hybrid systems, have been introduced by Davis in [Dav84, Dav93] for the finite dimensional setting and generalized in [BR11] for the infinite dimensional case. Recently, the asymptotic behavior of finite dimensional PDMPs has been investigated in [BLBMZ12, BLMZ12, TK09] through the research of an invariant measure and its uniqueness. Let us mention that, also for finite dimensional PDMPs, control problems have been studied in [CD10, CD11, Gor12], law of large numbers in [CDMR12, PTW10], numerical methods in [Rie12a, BDSD12], time reversal in [LP13], averaging in [FGC08, PTW12, WTP12] and to end up this list with no claim of completeness, estimation of the jump rates for PDMPs in [DHKR12, ADGP12b, Aza12]. As already mentioned, limit theorems for infinite dimensional PDMPs have been derived in [Aus08, RB13, RTW12, RT13]: a law of large numbers and Central Limit Theorems for sequences of Hilbert-valued PDMPs are obtained. Let us notice that a point process approach to PDMPs, not developed in the present thesis, is described in [Jac05]. The mathematical description of switching PDEs in the framework of PDMPs, as well as the presentation of their main properties and characteristics of these processes, is the purpose of the present paragraph.

Let $\mathcal{R}$ be a finite set and H a separable Hilbert space. The process we proceed to define has two distinct components: a continuous one with values in H and a jumping one with values in $\mathcal{R}$ which is *càdlàg* (right continuous with left limits).

For any $r \in \mathcal{R}$, let us consider $\Phi^r = (\Phi^r_t, t \in \mathbb{R}_+)$ a continuous dynamical system on H continuous in time and space with the semigroup property as follows:

i) $\Phi^r_t : H \to H$ is continuous for all $t \in \mathbb{R}_+$.

ii) Φ^r is a one parameter semigroup:

   a) $\Phi^r_0 = \mathrm{Id}_H$,
   b) $\Phi^r_{t+s} = \Phi^r_t \Phi^r_s$, for any $t, s \in \mathbb{R}_+$.

iii) for any $u \in H$, the map $t \mapsto \Phi^r_t(u)$ is continuous on $\mathbb{R}_+$.

This dynamical system describes the motion of the piecewise deterministic process between consecutive jumps. In the present work, these dynamical systems are defined as one parameter semigroups associated to well-posed PDEs.

We now define the jump mechanisms. For $r, \tilde{r} \in \mathcal{R}$, we define the jump rate functions $q_{r\tilde{r}} : H \to \mathbb{R}_+$ such that:

- $q_{r\tilde{r}}(u) \geq 0$ for $r \neq \tilde{r}$ and $q_{rr}(u) = 0$, for any $u \in H$.
- The mapping $q_{r\tilde{r}}$ is continuous on H.

Roughly speaking, if u were held fixed, the motion of the jumping component of the Piecewise Deterministic Process (PDP) would follow the dynamic of a time homogeneous continuous time Markov chain with generator $(q_{r\tilde{r}}(u))_{(r,\tilde{r})\in\mathcal{R}\times\mathcal{R}}$. That is, $q_{r\tilde{r}}$ would be the rate of jump from state r to state $\tilde{r}$ for the switching component of our PDP. We write $q_r(u)$ for the total rate of leaving the state r, namely $q_r(u) = \sum_{\tilde r\in\mathcal{R}\setminus\{r\}} q_{r\tilde r}(u)$ and define a family of survival functions for $(u,r)\in H\times\mathcal{R}$ by

$$S_{s,(u,r)}(t) = \exp\left(-\int_s^t q_r(\Phi^r_\tau(u))\mathrm{d}\tau\right)$$

for $s,t\in\mathbb{R}_+$ and a family of probability laws on $\mathcal{R}$ by

$$J_{(u,r)}(\tilde r) = \frac{q_{r\tilde r}(u)}{q_r(u)}.$$

Remark that $S_{s,(u,r)}(t)$ is well defined since the application $\tau\mapsto q_r(\Phi^r_\tau(u))$ is continuous, hence integrable, on $[s,t]$. To ensure the non explosion of the jump part of our PDMP in finite time, we assume that the total rates of leaving a state are uniformly bounded

$$\sup_{r\in\mathcal{R}}\sup_{u\in H} q_r(u) = q^+ < \infty. \tag{1.10}$$

The model defined by equations (1.4) and (1.5) is a PDMP in the present sense. Remember that E is a finite set, $I=[-l,l]$ and $\mathcal{N}=\mathbb{Z}\cap N(-l,l)$ for $N\geq 1$ and $l>0$. Let $H=H^1_0(I)$ denote the usual Sobolev space of functions in $L^2(I)$ with first derivative in the sense of distributions also in $L^2(I)$ and trace equal to zero on the boundary of I (c.f. Section 2.1). For $r\in\mathcal{R}=E^{\mathcal{N}}$ and $u_0\in H$, $\Phi^r_t(u_0)$ is the solution at time t of the PDE

$$\partial_t u_t = \partial_{xx}u_t + \frac{1}{N}\sum_{i\in\mathcal{N}} c_{r(i)}(v_{r(i)} - u_t(z_i))\delta_{z_i},$$

starting from u_0 at time 0 and with zero Dirichlet boundary conditions. One can shows that Φ^r is indeed a continuous dynamical system in H. In this case, the total rate of leaving a state $r\in\mathcal{R}$ is given by

$$q_r(u) = \sum_{i\in\mathcal{N}}\sum_{\xi\neq r(i)} \alpha_{r(i)\xi}(u(z_i))$$

for $u\in H$. Then, for $\tilde r$ which differs from r only by the component $r(i_0)$, the state $\tilde r$ is reached with probability

$$J_{(u,r)}(\tilde r) = \frac{\alpha_{r(i_0)\tilde r(i_0)}(u(z_{i_0}))}{q_r(u)}.$$

That is to say $q_{r\tilde r}(u) = \alpha_{r(i_0)\tilde r(i_0)}(u(z_{i_0}))$. If $\tilde r$ differs from r by two or more components, then $J_{(u,r)}(\tilde r) = 0$.

We now present the classical construction procedure for PDMPs considered in the present text. It is based on the original construction by Davis [Dav84].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space consisting in all sequences of independent uniformly distributed random variables on $[0,1]$. We construct the PDMP $(X_t, t \in \mathbb{R}_+) = ((u_t, r_t), t \in \mathbb{R}_+)$ from one such sequence.

We construct a PDMP $(u_t, r_t)_{t\geq 0}$ taking values in $H \times \mathcal{R}$. Let ν be a probability measure on $H \times \mathcal{R}$ and $(U_k, k \geq 0)$ be a sequence of independent random variables uniformly distributed on $[0,1]$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$. The measure ν is the law of the initial state $X_0 = (u_0, r_0)$. By virtue of Lemma 2.1.1 of [Rie12b], there exists a measurable function $f_0 : [0,1] \to H \times \mathcal{R}$ such that the law of $f_0(U_0)$ equals ν. Let $\omega \in \Omega$.

1. The initial condition is defined by
$$(u_0(\omega), r_0(\omega)) = f_0(U_0(\omega)).$$
2. The component $u(\omega)$ follows the deterministic dynamic given by the dynamical system $\Phi^{r_0(\omega)}$ as long as $r(\omega)$ remains equal to $r_0(\omega)$. The first jump time is defined by
$$T_1(\omega) = \inf\{t \geq 0; S_{0,(u_0(\omega), r_0(\omega))}(t) \geq U_1(\omega)\}. \tag{1.11}$$
Thus on $[0, T_1(\omega)[$ we have:
$$\left\{ \begin{array}{lcl} u_s(\omega) & = & \Phi_s^{r_0(\omega)}(u_0(\omega)), \\ r_s(\omega) & = & r_0(\omega). \end{array} \right.$$
In other words, equation (1.11) characterizes the law of T_1 conditional on (u_0, r_0) by its survival function.
3. At time $T_1(\omega)$, u is at $\Phi_{T_1}^{r_0}(u_0)$ since this component is continuous. The jumping component $r(\omega)$ is updated according to the law $J_{(u_{T_1(\omega)}(\omega), r_0(\omega))}$. There exists a measurable function $f_1 : [0,1] \to \mathcal{R}$ such that the law of $f_1(U_1)$ equals $J_{(u_{T_1(\omega)}(\omega), r_0(\omega))}$. Then
$$r_{T_1(\omega)}(\omega) = f_1(U_1(\omega)).$$
4. The whole process is obtained by re-iterating this algorithm. Let us give for example the second jump time which is defined by
$$T_2(\omega) = \inf\{t \geq T_1(\omega); S_{T_1(\omega),(u_{T_1(\omega)}(\omega), r_{T_1(\omega)}(\omega))}(T_1(\omega) + t) \geq U_2(\omega)\}. \tag{1.12}$$
Thus on $[T_1(\omega), T_2(\omega)[$:
$$\left\{ \begin{array}{lcl} u_s(\omega) & = & \Phi_s^{r_{T_1(\omega)}}(u_{T_1(\omega)}(\omega)), \\ r_s(\omega) & = & r_{T_1(\omega)}(\omega). \end{array} \right.$$

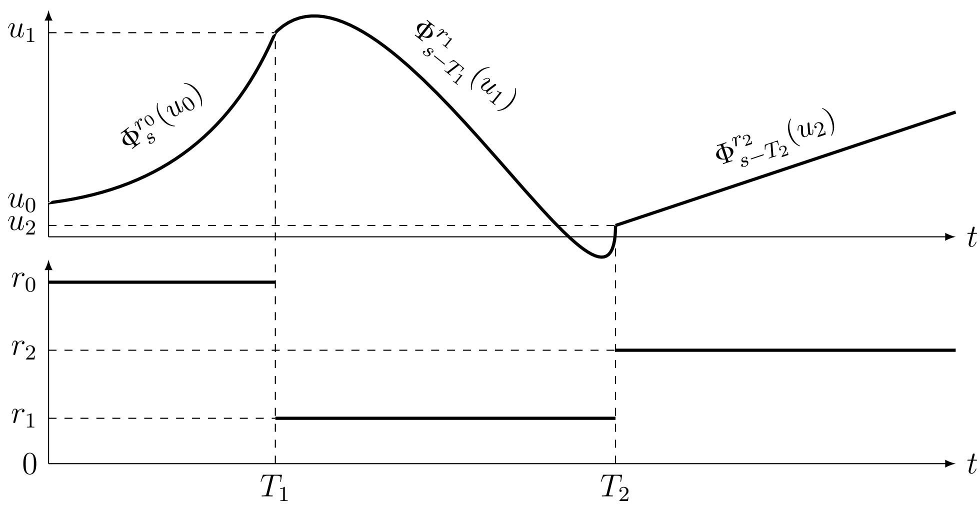


Figure 1.6: Typical behavior of the studied class of PDMPs.

From the above construction, it is clear that PDMPs are càdlàg piecewise deterministic process. The following theorem gathers the main properties of the presented class of PDMPs.

Theorem 1.2.1. *The stochastic process $X = ((u_t, r_t), t \in \mathbb{R}_+)$ defined above on $(\Omega, \mathcal{F}, \mathbb{P})$ enjoys the following properties:*

i) X is a càdlàg piecewise deterministic process.

ii) X is a strong *Markov process.*

iii) The domain $\mathcal{D}(\mathcal{A})$ of the extended generator $\mathcal{A}$ of X consists in all bounded measurable functions $\phi : H \times \mathcal{R} \to \mathbb{R}$ such that the map $t \mapsto \phi(\Phi_t^r(u), r)$ is absolutely-continuous for almost every $t \in \mathbb{R}_+$ (with respect to the Lebesgue measure on $\mathbb{R}_+$) for any $(u, r) \in H \times \mathcal{R}$.

iv) For $\phi \in \mathcal{D}(\mathcal{A})$, the extended generator is given by

$$\mathcal{A}\phi(u, r) = \frac{\mathrm{d}\phi}{\mathrm{d}t}(\Phi_t^r(u), r) + \sum_{\tilde{r} \in \mathcal{R}} q_{r\tilde{r}}(u)[\phi(u, \tilde{r}) - \phi(u, r)].$$

v) X is a Feller *process.*

The Markov property follows essentially from the semi-group property of the continuous dynamical systems $\{\Phi^r, r \in \mathcal{R}\}$ and the functional properties of the survival functions $S_{s,(r,u)}(t)$. The strong Markov property is proved in [Rie12b] Section 2.3.2 in our setting with arguments inspired by [Jac05], Theorems 7.3.2 and 7.5.1. As mentioned above, for the class of PDMPs we are considering there is no boundary in the state space for the continuous variable. Thus, there is no deterministic jumps. Moreover, we assume condition (1.10) which implies that there is no blow up in finite time. These two facts simplify the characterization of the domain of the extended generator, see [BLBMZ12, Rie12b]. We may notice that, as it can be expected, the generator splits itself into two parts: the first term describes the infinitesimal movement of the process between jumps, that is its deterministic behavior, whereas the second term describes at which rates the process jumps, that is the evolution of the jump component. A very straightforward proof of the Feller property may be derived from [BLBMZ12], Proposition 2.2. One could ask the question of the strong Feller property for the considered class of PDMPs. The strong Feller property is in general hard to prove, and often 'rare' for stochastic processes in infinite dimensions (see the note [Hai08]) except for Hilbert-valued Gaussian processes. Moreover, even for finite dimensional PDMPs, the strong Feller property fails in most cases.

For more details, we refer the reader to [BLBMZ12, Dav84, Rie12b].

1.2.2 A class of Stochastic Partial Differential Equations

The phenomenological models presented in Section 1.1.4 correspond mathematically to evolution equations and more precisely to reaction-diffusion partial differential equations. When adding noise to these models, they fall into the class of Stochastic Partial Differential Equations (SPDEs). In this thesis, we will consider system of SPDEs of the following form

$$\left\{ \begin{array}{rcl} \mathrm{d}u_t &=& [\nu\partial_{xx}u_t + f(u_t, w_t)]\mathrm{d}t + \sigma \mathrm{d}W_t, \\ \mathrm{d}w_t &=& g(u_t, w_t)\mathrm{d}t, \end{array} \right. \tag{1.13}$$

where ν is a constant of deterministic diffusion whereas σ is the intensity of the noise which can be seen as the strength of the stochastic diffusion. The (non linear) functions f and g are called the *reaction terms*. They describe the local dynamic of (1.13), that is the dynamic without the spatial diffusion ∂_{xx}. We always consider in this text that we are working with SPDEs on a bounded spatial domain D of $\mathbb{R}^d$. Most of the time, we will take d equal to 1 or 2, but this is not necessary. For us, $(W_t, t \geq 0)$ is a centered Gaussian process which can be white or colored in space. We now briefly recall how the noise is constructed. Then we list its principal properties and we state the well-definiteness of a solution to (1.13).

As in Section 1.2.1, H denotes a real separable Hilbert space with scalar product $(\cdot,\cdot)$ and associated Hilbert basis $\{e_k, k \geq 1\}$. Let Q be a non negative symmetric linear operator on H of trace class. This means that

- $\forall \phi \in H, \quad (Q\phi, \phi) \geq 0$ (non negativity);
- $\forall (\phi_1, \phi_2) \in H \times H, \quad (Q\phi_1, \phi_2) = (\phi_1, Q\phi_2)$ (symmetry);
- $\mathrm{Tr}(Q) = \sum_{k\geq 1} (Qe_k, e_k) < \infty$ (trace class).

Actually, in our definition, Q is symmetric everywhere so that Q is a self-adjoint operator.

Definition 1.2.1 (Q-Wiener process.)**.** *There exists a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *on which we can define a stochastic process* $(W_t^Q, t \in \mathbb{R}_+)$ *such that*

i) for each $t \in \mathbb{R}_+$, W_t^Q *is an* H*-valued random variable.*

ii) $(W_t^Q, t \in \mathbb{R}_+)$ *is a Lévy process, that is, it is a process with independent and stationary increments:*

a) for any sequence $t_1, \dots, t_n$ *of strictly increasing times the random variables* $W_{t_2}^Q - W_{t_1}^Q, \dots, W_{t_n}^Q - W_{t_{n-1}}^Q$ *are independent (independent increments).*

b) for any two times $s < t$ *the random variable* $W_t^Q - W_s^Q$ *has the same law as* W_{t-s}^Q *(stationary increments).*

iii) for any $t \in \mathbb{R}_+$ *and any* $\phi \in H$, (W_t^Q, ϕ) *is a real centered Gaussian variable with variance* $t(Q\phi, \phi)$*.*

iv) $(W_t^Q, t \in \mathbb{R}_+)$ *is continuous in time* $\mathbb{P}$*-almost surely.*

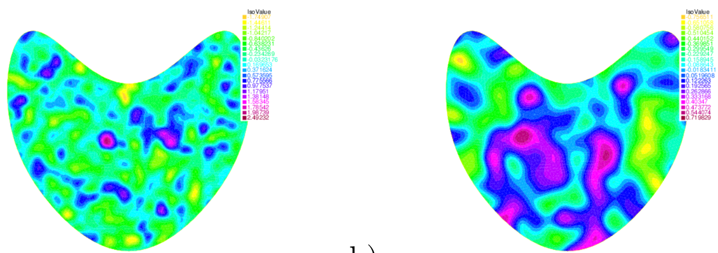
The stochastic process $(W_t^Q, t \in \mathbb{R}_+)$ may be built in the following way. Since Q is a non negative self adjoint operator of trace class, we can consider its square root $Q^{\frac{1}{2}}$. This is an operator on H with adjoint denoted by $(Q^{\frac{1}{2}})^*$ and such that $Q = (Q^{\frac{1}{2}})^* Q^{\frac{1}{2}}$. Let $(\beta_k, k \geq 1)$ be a sequence of independent real Brownian motions defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The sequence of processes defined by

$$W_t^{Q,n} = \sum_{k=1}^{n} \beta_{k,t} Q^{\frac{1}{2}} e_k,$$

converges in $L^2(\Omega, H)$ towards a process W^Q satisfying the conditions of Definition 1.2.1. The fact that Q is of trace class keeps the things rather easily to understand. We say that the noise W^Q is *colored* by Q. However, one may consider symmetric non-negative operators which are not of trace class, the most common choice being the identity operator. In this case, Definition 1.2.1 makes sense if one enlarges the Hilbert space H. This defines what is generally called a *white* noise which we denote simply by $(W_t, t \in \mathbb{R}_+)$. For white noise the construction

$$W_t = \sum_{k \geq 1} \beta_{k,t} e_k,$$

is formal but can be made rigorous in a Hilbert space containing H.



a) b)

Figure 1.7: Simulation of two different colored noises on a bounded domain D of $\mathbb{R}^2$ with covariance operator on $L^2(D)$ given by $Q\phi(x) = \int_D \phi(y)(4\xi^2)^{-1} e^{-\frac{\pi^2}{4\xi^2}|x-y|^2} \mathrm{d}y$ for $x \in D$. a) $\xi = 1$, small spatial correlations. b) $\xi = 2$, higher range spatial correlations.

In numerous places in the text we will encounter stochastic evolution equations of the form

$$\mathrm{d}X = [AX + F(X)]\mathrm{d}t + \mathrm{d}W^Q, \tag{1.14}$$

where the noise W^Q is written in the form $Q^{\frac{1}{2}}\mathrm{d}W$ where $(W_t, t \in \mathbb{R}_+)$ is a white noise. We explain briefly how we can show equation (1.14) has a unique solution. The noise is additive, that is, it does not depend on X and is in some sense simply added to the deterministic evolution equation

$$dX = [AX + F(X)]\mathrm{d}t. \tag{1.15}$$

In an integral form (the *mild* form in the stochastic context) a solution of (1.14) can be written as

$$X_t = e^{At}X_0 + \int_0^t e^{A(t-s)} F(X_s)\mathrm{d}s + \int_0^t e^{A(t-s)} \mathrm{d}W_s^Q. \tag{1.16}$$

The trick is to consider the stochastic convolution $\int_0^t e^{A(t-s)}\mathrm{d}W_s^Q$ as a *control* term. Let us replace the stochastic convolution in (1.16) by a control $(\omega_t, t \in \mathbb{R}_+)$ having the same properties of regularity in time and space as the stochastic convolution. If the deterministic equation with control ω, namely

$$X_t = e^{At}X_0 + \int_0^t e^{A(t-s)}F(X_s)\mathrm{d}s + \omega_t, \tag{1.17}$$

has a unique solution, then the corresponding stochastic equation (1.16) will have a unique solution too. One may call a measurable selection theorem to prove this fact. These kinds of theorems allow to select in an appropriate way the stochastic convolution $\int_0^t e^{A(t-s)}\mathrm{d}W_s^Q$ inside the class of controls ω, see [Wag80].

We end this section with some words about the martingale problem associated to the SPDE (1.14). The formalism is taken from [Zam00]. Under usual hypotheses on the operators A and F such that A generates a strongly continuous semi-group $(e^{At}, t \in \mathbb{R}_+)$ satisfying $\|e^{At}\| \leq 1$ and F is θ-Hölder continuous for some $\theta \in (0,1)$, one can consider the infinitesimal generator $\mathcal{K}$ associated to $(X_t, t \in \mathbb{R}_+)$ defined by

$$\forall \phi \in \mathcal{D}(\mathcal{K}), \quad \forall u \in H, \quad \mathcal{K}\phi(u) = \frac{\mathrm{d}\phi}{\mathrm{d}u}(u)[Au + F(u)] + \frac{1}{2}\mathrm{Tr}\left(Q\frac{\mathrm{d}^2\phi}{\mathrm{d}u^2}(u)\right).$$

Under these assumptions, $\mathcal{D}(\mathcal{K})$, the domain of $\mathcal{K}$, is a subset of $\mathcal{C}^{2+\theta}(H)$. Given a time horizon T, let us define the coordinate process $(X_t, t \in [0,T])$ on $H^{[0,T]}$ by

$$X_t : \omega \in H^{[0,T]} \mapsto \omega(t) := X_t(\omega) \in H.$$

Definition 1.2.2 (Martingale problem for SPDEs)**.** *A solution to the martingale problem associated to (1.14) with $X_0 = x \in H$ (x held fixed) is a probability measure $\mathbb{P}_x$ on $H^{[0,T]}$ such that*

$$\mathbb{P}_x\left(\omega \in H^{[0,T]} | \omega(0) = x, \omega \textit{ is Borel}\right) = 1$$

and the process

$$\left(\phi(X_t) - \int_0^t \mathcal{K}\phi(X_s)\mathrm{d}s, t \in \mathbb{R}_+\right)$$

is a $\mathbb{P}_x$-martingale for all $\phi \in \mathcal{D}(\mathcal{K})$.

For more details, we refer the reader to [DPZ92, DP04, PZ07] which are classical references for SPDEs.

1.3 Main results of the thesis

This section displays the main results of the present thesis chapter by chapter, except for Chapter 2 which is devoted to mathematical preliminaries.

1.3.1 Chapter 2: mathematical preliminaries.

Chapter 2 gathers mathematical preliminaries used in the proof of the main theorems in Chapters 3, 4 and 5. After reviewing basic facts of functional analysis on Hilbert spaces (Section 2.1), we prove some useful estimates about the heat semigroup on a segment with zero Dirichlet boundary conditions in Section 2.2. We recall a tightness criterion for Hilbert-valued stochastic processes that was proved in [Mé84] (c.f. Section 2.3).

1.3.2 Chapter 3 and 4: averaging for a class of multiscale conductance-based models.

In **Chapters 3** and **4**, we consider a general class of spatially extended stochastic conductance-based models with a finite number of ion channels like those introduced in Section 1.1.3 but moreover with multiple timescales. We first recall that these models are described mathematically as Hilbert-valued PDMPs, c.f. Section 1.2.1. In line with results obtained for finite dimensional conductance-based models [PTW12, WTP12], that is without spatial extension, we obtain when a timescale ratio ε goes to zero, the so-called averaged model. We introduce different timescales in spatially extended conductance-based models considering that some ion channels open and close at faster rates than others. We perform a slow-fast analysis of this model and prove that asymptotically, it reduces to the averaged model which is still a PDMP in infinite dimensions for which we provide effective evolution equations and jump rates. We thus reduce the complexity of the original model by simplifying the kinetics of ion channels. The natural step after this averaging result is to prove the Central Limit Theorem associated to it. Thus, we further study the fluctuations around the averaged system in the form of a Central Limit Theorem. We apply the mathematical results of averaging and fluctuations to the Hodgkin-Huxley and Morris-Lecar models with stochastic ion channels. To the best of our knowledge, no averaging results were available for Hilbert-valued PDMPs. For PDMPs in finite dimension, a theory of averaging has been developed in [FGC08] and central limit theorem has been proved in [PTW12].

Namely, we consider the process $((u^\varepsilon_t, r^\varepsilon_t), t \in [0,T])$, $T > 0$, described by equations (1.4) and (1.5) in we introduce in these equations a timescale separation parameter $\varepsilon > 0$. More precisely, the evolution equation satisfied by u^ε reads

$$\partial_t u^\varepsilon_t = \partial_{xx} u^\varepsilon_t + \frac{1}{N}\sum_{i\in\mathcal{N}} c_{r_t(i)}(v_{r_t(i)} - u^\varepsilon_t(z_i))\delta_{z_i}, \tag{1.18}$$

and the jump component r^ε satisfies

$$\mathbb{P}(r^\varepsilon_{t+h}(i) = \zeta | r^\varepsilon_t(i) = \xi) = \frac{1}{\varepsilon}\alpha_{\xi\zeta}(u^\varepsilon_t(z_i))h + \mathrm{o}(h) \tag{1.19}$$

for $i \in \mathcal{N}$ and $\zeta \neq \xi$ in E. Recall that the axon is represented as a segment I and that a configuration for the ion channels is an element r of the set $\mathcal{R} = E^{|\mathcal{N}|}$. The quantity $|\mathcal{N}| = N$ represents the number of ion channels and E is the state space of each of them. We show in Proposition 3.2.1 and 3.2.2 that $((u_t^\varepsilon, r_t^\varepsilon), t \in [0,T])$ is an almost-surely uniformly bounded in ε stochastic process valued in $\mathcal{C}([0,T], H_0^1(I)) \times \mathbb{D}([0,T], \mathcal{R})$. We recall that $\mathbb{D}([0,T], \mathcal{R})$ stands for the space of $\mathcal{R}$-valued càdlàg functions on $[0,T]$ endowed with the Skorohod topology. The infinitesimal generator $\mathcal{A}^\varepsilon$ of the stochastic process $(u^\varepsilon, r^\varepsilon)$ is defined, for a Fréchet differentiable (w.r.t. u) function f in the domain $\mathcal{D}(\mathcal{A}^\varepsilon)$ of $\mathcal{A}^\varepsilon$ such that $\frac{\mathrm{d}f}{\mathrm{d}u}(u,r)$ is in V for $(u,r) \in V \times \mathcal{R}$, by

$$\mathcal{A}^\varepsilon f(u,r) = \frac{\mathrm{d}f}{\mathrm{d}u}(u,r)[\partial_{xx} u + G_r(u)] + \mathcal{B}^\varepsilon f(u,\cdot)(r), \tag{1.20}$$

where $G_r(u) = \frac{1}{N}\sum_{i\in\mathcal{N}} c_{r(i)}(v_{r(i)} - u(z_i))\delta_{z_i}$. The quantities $\frac{\mathrm{d}f}{\mathrm{d}u}$ and $f(u,\cdot)$ denote respectively the Fréchet derivative of f with respect to u (see Section 2.4 for more details) and the vector $(f(u,r), r \in \mathcal{R})$. Since $\mathcal{B}^\varepsilon$, the jump part of the generator, is proportional to ε (c.f. (1.19)), the process has two distinct timescales. The variable r^ε jumps on a fast timescale between the states of $\mathcal{R}$ according to the jump's dynamic (1.19) and between the jumps, the variable u^ε evolves on a slow timescale according to the evolution equation (1.18). The above setting with $\mathcal{B}^\varepsilon = \frac{1}{\varepsilon}\mathcal{B}$ will be called "all-fast", in contrast with a more general case where the jumping part of the generator has an expansion in a slow and a fast part

$$\mathcal{B}^\varepsilon = \frac{1}{\varepsilon}\mathcal{B} + \hat{\mathcal{B}},$$

where $\mathcal{B}$ is the fast part of the generator, gathering the states of $\mathcal{R}$ communicating at rate of order $\frac{1}{\varepsilon}$ and $\hat{\mathcal{B}}$ is the slow part of the generator, taking account of the communications happening at rate of order 1. This case is referred to as the "multiscale" case since $\mathcal{B}^\varepsilon$ decomposes itself in a slow and fast part. For simplicity of presentation in this introduction, we state the results of Chapters 3 and 4 in the all-fast case but emphasize that the corresponding results are stated and proved in the multiscale case in Chapters 3 and 4.

For any fixed $u \in H_0^1(I)$ the generator $\mathcal{B}$ has a unique stationary distribution on $\mathcal{R}$ denoted by $\mu(u) = \bigotimes_{i\in\mathcal{N}} \mu(u(z_i))$ where for $i \in \mathcal{N}$, $\mu(u(z_i))$ is a measure on E related to the stationary measure associated to the dynamic of the ionic channel at location z_i. Then we define the averaged reaction term by

$$\begin{aligned} F(u) &= \int_{\mathcal{R}} G_r(u)\mu(u)(\mathrm{d}r) \\ &= \frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}} c_\xi \mu(u(\frac{i}{N}))(\xi)(v_\xi - u(\frac{i}{N}))\delta_{\frac{i}{N}}. \end{aligned} \tag{1.21}$$

In the all-fast case, the averaged equation reads

$$\partial_t u = \Delta u + F(u) \tag{1.22}$$

We have the following result.

Theorem 1.3.1. *The stochastic process u^ε solution of (1.18-1.19) converges in law to a solution of (1.22) when ε goes to* 0.

Theorem 1.3.1 is the main result of Chapter 3. Of course, in the multiscale case, the averaged model is no longer deterministic since jumps with frequency of order 1 remain and the averaged limit is no longer deterministic but a PDMP. As an application, we consider a stochastic version of the Hodgkin-Huxley model with no potassium and leak currents: we only consider the sodium current. This is still a case of interest since it is the sodium channels that are involved in the generation of a spike through the inward current. Numerical simulations are provided. They show the tractability of our approach and the consistency of averaged stochastic model with the averaged version of the corresponding deterministic Hodgkin-Huxley equations.

We then proceed to analyze the fluctuations around the averaged limit when the Dirac mass of the model δ_{z_i} are replaced by the mollifiers ϕ_{z_i}. The results stated above (Theorem 1.3.1) remains valid with mollifiers with the appropriate modifications (for example, all the occurrences of $u(z_i)$ are replaced by $(u, \phi_{z_i})_{L^2(I)}$). To obtain a non trivial result, the difference between the non-averaged and averaged processes must be renormalized. Let us define for this purpose the stochastic process

$$z^\varepsilon = \frac{u^\varepsilon - u}{\sqrt{\varepsilon}}$$

for $\varepsilon > 0$. We denote by $\{f_k, k \geq 1\}$ a Hilbert basis of $L^2(I)$.

Theorem 1.3.2. *When ε goes to 0 the process z^ε converges in distribution in $\mathcal{C}([0,T], L^2(I))$ towards a process z. For $u \in L^2(I)$, let $C(u) : L^2(I) \to L^2(I)$ be a diffusion operator characterized by*

$$(C(u)f_j, f_i)_{L^2(I)} \quad = \quad \int_{\mathcal{R}} (G_r(u) - F(u), f_i)_{L^2(I)} (\Phi(r,u), f_j)_{L^2(I)} \mu(u)(\mathrm{d}r),$$

Φ is the unique solution of the equation

$$\begin{cases} \mathcal{B}(u)\Phi(r,u) & = & -(G_r(u) - F(u)) \\ \int_{\mathcal{R}} \Phi(r,u)\mu(u)(\mathrm{d}r) & = & 0. \end{cases}$$

Let us also define an operator $\bar{\mathcal{G}}^1$ *by*

$$\bar{\mathcal{G}}^1(t)\psi(z) = \frac{\mathrm{d}\psi}{\mathrm{d}z}(z)\left[\Delta z + \frac{\mathrm{d}F}{\mathrm{d}u}(u_t)[z]\right] + \mathrm{Tr}\left[\frac{\mathrm{d}^2\psi}{\mathrm{d}z^2}(z)C(u_t)\right],$$

for $t \in [0,T]$ *and a measurable, bounded and twice Fréchet differentiable function* $\psi : L^2(I) \to \mathbb{R}$. *The process* z *is uniquely determined as the solution of the following martingale problem. For any measurable, bounded and twice Fréchet differentiable function* $\psi : L^2(I) \to \mathbb{R}$, *the process*

$$\bar{N}_\psi(t) := \psi(z_t) - \int_0^t \bar{\mathcal{G}}^1(s)\psi(z_s)\mathrm{d}s$$

for $t \in [0,T]$, *is a martingale.*

The evolution equation associated to the martingale problem stated in Theorem 1.3.2 is the following stochastic partial differential equation

$$\mathrm{d}z_t = \left(\Delta z_t + \frac{\mathrm{d}F}{\mathrm{d}u}(u_t)[z_t]\right)\mathrm{d}t + \Gamma(u_t)\mathrm{d}W_t. \tag{1.23}$$

In particular, the limiting process z solves an SPDE. In the multiscale case, the corresponding limiting process is an hybrid SPDE since the jumps occurring at order 1 remain in the limit. As an application, we consider the stochastic Morris-Lecar model where all the quantities of interest may be computed analytically and simulated numerically. For example, the function $u \mapsto \mathrm{Tr}\ C(u)$ may be computed and provides a quantitative measure of the variations of u^ε around u.

The proof of Theorem 1.3.1 is based on the Prohorov method: the existence of a limit in law for a sequence of stochastic processes is equivalent to the tightness of this sequence. At this point, we use the tightness criterion in Hilbert spaces recalled in Section 2.3. Then, the law of the limiting points must be identified or characterized. For Theorem 1.3.1, we use the weak form of the evolution equation (1.18) to identify the limit.

Theorem 1.3.2 is proved by tightness and identification of the limit like Theorem 1.3.1. But this time, in order to identify of the limit, we develop the infinitesimal generator of the process $(z^\varepsilon, r^\varepsilon)$ in power of ε in order to identify its first order terms.

1.3.3 Chapter 5: quantitative ergodicity for infinite dimensional PDMPs.

In **Chapter 5**, our aim is to average Hilbert-valued PDMPs when continuous component and some jumping component both evolve fast. In Chapters 3 and 4, in order to derive averaging results, it was of first importance to get informations on the invariant measure associated to the fast motion of the ionic channels. Indeed, it was thanks to these informations on the invariant measure that we were able to prove that the slow-fast PDMP converges to the averaged limit when accelerating some of the jumps of the switching process. Since now the fast part of our PDMP is itself a PDMP, and not only a continuous time Markov chain as in Chapters 3 and 4, we start with investigating the existence of the invariant measure for Hilbert-valued PDMPs and the rate of convergence towards this measure.

Let E be a finite set and H a separable Hilbert space with Hilbert basis $\{e_k, k \geq 1\}$. On one hand, for any fixed $i \in E$ we consider the non-linear problem

$$\partial_t u_t = A_i u_t + F_i(u_t) \tag{1.24}$$

with $u_0 \in H$ for $t \in \mathbb{R}_+$. The linear and non-linear operators A_i and F_i are either dissipative or Lipschitz such that the system (5.1) has a unique solution u which belongs to $\mathcal{C}([0,T],H)$ for any time horizon $T \geq 0$. On the other hand, for any fixed $u \in H$, we consider an E-valued continuous time Markov chain $I = (I_t, t \geq 0)$ with u-dependent generator $Q(u) = (q_{ij}(u))_{i,j}$ with some boundedness conditions on the functions q_{ij}.
In line with recent results of [BLBMZ12] for finite dimensional PDMPs, Chapter 5 is concerned with the long time behavior of PDMPs of the form

$$\left\{\begin{array}{rcl} \partial_t u_t & = & A_{I_t} u_t + F_{I_t}(u_t), \\ \mathbb{P}(I_{t+h} = j | I_t = i) & = & q_{ij}(u_t)h + \mathrm{o}(h), \quad i \neq j \end{array}\right. \tag{1.25}$$

for $t \geq 0$ and given random initial conditions $u_0 \in H$ and $I_0 \in E$. The process $(u_t, I_t)_{t\geq 0}$ is a well-defined PDMP. We prove existence and uniqueness of the invariant measure and the convergence of u toward this measure in Wasserstein distance. See Chapter 5 for the definition of Wasserstein distance $\mathcal{W}$ and its main properties. Let us just say that this convergence is equivalent to convergence in law *plus* convergence of moments of order 1 or higher depending of the order of the Wasserstein distance.

Proposition 1.3.1. *Let $\mathcal{L}(u_t)$ denote the law of u at time t. The process $(u_t, t \geq 0)$ has a unique invariant measure ν on H such that*

$$\mathcal{W}(\mathcal{L}(u_t), \nu) \leq \alpha(1+t)e^{-\beta t},$$

where α and β are two positive constants which we know explicitly.

That is, the rate of convergence toward the invariant measure is exponential. Proposition 1.3.1 is proved thanks to coupling arguments developed in [BLBMZ12] and extended to our infinite dimensional framework. Moreover we show the convergence of the invariant measure of the truncation of u up to the order N toward the invariant measure of u under the Wasserstein metric.

Proposition 1.3.2. *For $N \in \mathbb{N}$, let $u^{(N)}$ be the truncation of u up to the order N: $u_t^{(N)} = \sum_{k=1}^N (u_t, e_k)e_k$ for $t \geq 0$. Then $u^{(N)}$ has a unique invariant measure $\nu^{(N)}$ which converges toward ν when N goes to infinity. More precisely, if $\mathbb{E}(\|u_0\|^2)$ is finite we have*

$$\mathcal{W}(\nu^{(N)}, \nu) \leq \sqrt{a_N}$$

for any $N \in \mathbb{N}$ where the sequence $(a_N)_{N\in\mathbb{N}}$ goes to zero when N goes to infinity and is explicit.

As an application, we consider the averaging of a fast PDMP fully coupled to a slow continuous time Markov chain, a situation motivated by the study of conductance-based neuron models. Indeed, the potential and some ionic species are often assumed to evolve at a faster timescale than other ionic species. For instance, in the Hodgkin-Huxley model, the potential and the gates of type m have a faster dynamic than the gates of type h and n, see the description of the Hodgkin-Huxley model in Section 1.1.2 and also [RW07] where averaging is considered in this situation for the deterministic Hodgkin-Huxley model. Thus, it is of first importance to analyze the behavior of such a system at the first order in ε, that is the averaging of the model.

Namely, this leads us to consider, a PDMP of the following form where $\varepsilon \in (0, 1)$

$$\left\{\begin{array}{rcl} \partial_t u_t &=& \frac{1}{\varepsilon}\left[A_{I_t^{(1)},I_t^{(2)}} u_t + F_{I_t^{(1)},I_t^{(2)}}(u_t)\right], \quad u_0 \in H, \\ \mathbb{P}(I_{t+h}^{(1)} = j | I_t^{(1)} = i) &=& \frac{1}{\varepsilon} q_{ij}^{(1)}(u_t)h + \mathrm{o}(h), \quad i \neq j, \quad i,j \in E^{(1)}, \\ \mathbb{P}(I_{t+h}^{(2)} = j | I_t^{(2)} = i) &=& q_{ij}^{(2)}(u_t)h + \mathrm{o}(h), \quad i \neq j, \quad i,j \in E^{(2)} \end{array}\right. \tag{1.26}$$

for $t \geq 0$. The PDMP $(u, I^{(1)}, I^{(2)})$ gives rise to two distinct dynamic. The PDMP $(u, I^{(1)})$ evolves faster than the process $I^{(2)}$ according to the timescale separation introduced by the small parameter ε. On a fixed time horizon $[0, T]$, when ε goes to zero, the process $(u, I^{(1)})$, denoted in the sequel by $(u^\varepsilon, I^{(1),\varepsilon})$, will rapidly reach its stationary behavior. Then the slow process $I^{(2)}$, denoted in the sequel by $I^{2,\varepsilon}$, will evolve according to the averaged dynamic of $(u, I^{(1)})$. The process $(u, I^{(1)})$ will be replaced by its invariant law.
Let us assume that in system (1.26), the process $I^{(2),\varepsilon}$ is frozen to the value $i^{(2)} \in E^{(2)}$. Then, applying Proposition 1.3.1 above, we know that the PDMP defined by

$$\left\{\begin{array}{rcl} \partial_t u_t &=& \left[A_{I_t^{(1)},i^{(2)}} u_t + F_{I_t^{(1)},i^{(2)}}(u_t)\right], \quad u_0 \in H, \\ \mathbb{P}(I_{t+h}^{(1)} = j | I_t^{(1)} = i) &=& q_{ij}^{(1)}(u_t)h + \mathrm{o}(h), \quad i \neq j, \quad i,j \in E^{(1)}, \end{array}\right. \tag{1.27}$$

has a unique invariant measure $\mu_{i^{(2)}}$ on $H \times E^{(1)}$. Let us average the dynamic of $I^{(2)}$ against the invariant measure of $(u, I^{(1)})$. For any $i, j \in E^{(2)}$ we define the averaged jump rate from i to j

$$\bar{q}_{ij} = \int_H q_{ij}^{(2)}(u)\nu_i(\mathrm{d}u). \tag{1.28}$$

We denote by $\bar{Q}$ the intensity matrix associated to the averaged jump rates $\bar{q}_{ij}$ and by $\bar{J} = (\bar{J}_t, t \in [0,T])$ the continuous time Markov chain associated to $\bar{Q}$ with for simplicity: $\bar{J}_0 = I_0^{(2),\varepsilon} = i_0^{(2)} \in E^{(2)}$. We obtain the following averaged result for the sequence of processes $(I^{(2),\varepsilon}, \varepsilon \in (0,1))$.

Theorem 1.3.3. *The process $I^{(2),\varepsilon} = (I_t^{(2),\varepsilon}, t \in [0,T])$ converges in law when ε goes to zero towards the continuous time Markov chain $\bar{J}$. Moreover the order of convergence is* 1 *in the sense that*

$$\sup_{t\in[0,T]} |\mathbb{E}(\phi(I_t^{(2),\varepsilon}) - \phi(\bar{J}_t))| = \mathrm{O}(\varepsilon)$$

for any real valued measurable and bounded function ϕ.

Theorem 1.3.3 implies that at first order in ε, the slow-fast system (1.26) reduces to a continuous time Markov chain. If we consider system (1.26) as a model for a neural cell, we notice that the equation on the potential has disappeared in the averaged model and the potential is only present through the invariant measure of the fast part of the system. This may look odd since the potential is a variable of first interest in conductance-based neuron models. As remarked in [RW07] for the finite dimensional deterministic Hodgkin-Huxley model, one may certainly consider an intermediate model between the three dimensional two timescales model (1.26) and the one dimensional averaged model following the dynamic of $\bar{J}$. That is a simplified model that still contains an equation on the potential remains.

1.3.4 Chapter 6: simulations of SPDEs for excitable media.

Chapter 6 is concerned with the numerical simulation of SPDEs used to model excitable cells in order to analyze the effect of noise on such biological systems. Our aim is twofold. The first is to propose an efficient and easy-to-implement method to simulate this kind of models. The second is to analyze the effect of noise on these systems thanks to numerical experiments. Namely, in models for cardiac cells, we investigate the possibility of purely noise induced reentrant patterns such as spiral or scroll-waves since these phenomena are related to major troubles of the cardiac rhythm such as tachyarrhythmia.

We will consider two phenomenological stochastic models derived from classical deterministic ones: the Barkley and the Mitchell-Schaeffer models. Mathematically, they consist in a system of PDEs driven by a colored noise. More precisely, they may be written

$$\left\{\begin{array}{lcl} \mathrm{d}u & = & [\nu\Delta u + \frac{1}{\varepsilon}f(u,v)]\mathrm{d}t + \sigma\mathrm{d}W^Q, \\ \mathrm{d}v & = & g(u,v)\mathrm{d}t, \end{array}\right. \tag{1.29}$$

on $[0,T] \times D$, where D is a regular bounded open set of $\mathbb{R}^2$ or $\mathbb{R}^3$. This system is completed with boundary and initial conditions. W^Q is a colored Gaussian noise source as defined in Section 1.2.2. The general structure of f and g is also typical of excitable dynamics. In particular, in the models that we will consider, the neutral curve $f(u,v) = 0$ when v is held fixed is cubic in shape.

To achieve our first aim, that is to numerically compute a solution of system (1.29), we work with a numerical scheme based on finite difference discretization in time and finite element method in space. The choice of finite element discretization in space has been directed by two considerations. The first one is that this method fits naturally to a general spatial domain: we want to investigate the behavior of solutions to (1.29) on domains with various geometries. The second one is that it allows to implement the numerical scheme using popular software such as the finite element software FreeFem++ or equivalent. The discretization of SPDEs by finite differences in time and finite elements in space has been considered by several authors in theoretical studies, see for example [DP09, Deb11, CYY07, LT12, Wal05]. Other methods of discretization are considered for example in [ANZ98, GM05, GMV12, Jen09, JR10, KLNS11, KLL10, LT10, Yan05]. These methods are based on finite difference discretization in time coupled either to finite difference in space or to the Galerkin spectral method, or to the finite element method on the integral formulation of the evolution equation. We emphasize that we do not consider a Galerkin spectral method or exponential integrator, that is, roughly speaking, we neither use the spectral decomposition of the solution of (1.29) according to a Hilbert basis of $L^2(D)$ (or an other Hilbert space related to D) nor the semigroup attached to the linear operator (the Laplacian in (1.29)), in order to build our scheme. We only use the variational version of the finite element method in order to fit to commonly used finite elements method for deterministic PDEs. Let us note that Chapter 6 is more numerically oriented than the above cited papers, in the spirit of [Sha05]. In [Sha05], the author numerically analyzes the effect of noise on excitable systems thanks to a Galerkin spectral method of discretization on the square. We pursue the same objective using the finite element method instead of the Galerkin spectral one. Let us notice that a discretization scheme for SPDEs driven by white noise for spatial domains of dimension greater or equal to 2 may lead to non trivial phenomena, see [HRW12]. Considering colored noises may also be seen as a way to circumvent these difficulties.

As is well known, one can consider two types of errors related to a numerical scheme for stochastic evolution equations: the strong error and the weak error.

The strong error for the discretization we consider has been analyzed for one dimensional spatial domains (line segments) in [Wal05]. The weak error for more general spatial domains, of dimension 2 or 3 for example, has been considered in [DP09, Deb11]. In Chapter 6, we prefer to consider the strong error of convergence of our scheme because we want to investigate numerally pathwise properties of the model. Working with spatial domains of dimension 2 or 3, we show that the strong order of convergence of the considered method for a class of linear stochastic parabolic equations is twice less than the weak order obtained in [DP09]. This is what is expected since this same duality between weak and strong order holds for the discretization of finite dimensional stochastic differential equations (SDEs).

Namely, let us consider the following linear parabolic stochastic equation on $[0,T]\times D$

$$\left\{\begin{array}{rcl} \mathrm{d}u_t & = & \Delta u_t \mathrm{d}t + \sigma \mathrm{d}W_t^Q, \\ u_0 & = & \xi. \end{array}\right. \tag{1.30}$$

The initial condition ξ is a $L^2(D)$-valued random variable. For simplicity in the introduction, the operator Δ is the Laplacian operator with domain $H^2(D)\cap H_0^1(D)$. More general parabolic linear operators are considered in Chapter 6.

Let $\mathcal{T}_h$ be a family of triangulation of the domain D by triangles ($d=2$) or tetrahedra ($d=3$). The size of $\mathcal{T}_h$ is given by

$$h = \max_{T\in\mathcal{T}_h} h(T),$$

where $h(T) = \max_{x,y\in T}|x-y|$ is the diameter of the element T. Let $\{P_i, 1\le i\le N_h\}$ be the set of all the nodes associated to the triangulation $\mathcal{T}_h$. The basis for the P1 finite element method is given by

$$\mathcal{B}_{1,\mathcal{T}_h} = \{\psi_i, 1\le i\le N_h\},$$

where ψ_i is the continuous piecewise affine function on D defined by $\psi_i(P_j)=\delta_{ij}$ (Kronecker symbol) for all $1\le i,j\le N_h$ if P_j is in the interior of D and $\psi_i(P_j)=0$ for all $1\le i,j\le N_h$ if P_j is on the boundary of D. We denote by V_h this space of P1 finite elements. The P1 approximation of the noise is

$$W_t^{Q,h,1} = \sum_{i=1}^{N_h} W_t^Q(P_i)\psi_i. \tag{1.31}$$

P0 finite elements are also considered in Chapter 6.

We study the following numerical scheme to approximate equation (1.30) defined recursively by: for u_0 given in V_h, find $(u_n^h)_{0\le n\le N}$ in V_h such that for all

$n \leq N-1$

$$\left\{ \begin{array}{rcl} \frac{1}{\Delta t}(u_{n+1}^h - u_n^h, v_h) + (\nabla u_{n+1}^h, \nabla v_h) & = & \frac{\sigma}{\Delta t}(W_{(n+1)\Delta t}^{Q,h,1} - W_{n\Delta t}^{Q,h,1}, v_h) \\ u_0^h & = & u_0 \end{array} \right. \tag{1.32}$$

for all $v_h \in V_h$.

Proposition 1.3.3. *There exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying the usual conditions such that if $\sqrt{\mathbb{E}(\|u_0^h - u_0\|^2)} = \mathrm{O}(h)$ then, for all $0 \leq n \leq N$ we have*

$$\sqrt{\mathbb{E}(\|u_n^h - u_{n\Delta t}\|^2)} = \mathrm{O}(h + \sqrt{\Delta t}), \tag{1.33}$$

where the estimate is uniform in $n \in \{0, \ldots, N\}$.

Thanks to the spatial regularity of the considered noise, the proof we provide follows classical arguments used to analyze the error introduced by the deterministic finite element method [RT83].

Our motivation for considering systems such as (1.29) comes from biological considerations. In the cardiac muscle, tachyarrhythmia are disturbances of the heart rhythm in which the heart beating rate is abnormally increased. This major trouble of the cardiac rhythm may lead to rapid loss of consciousness and to death. As explained in [Hin02, JC06], the vast majority of tachyarrhythmia are perpetuated by a reentrant mechanism. It is well known that deterministic excitable systems of type (1.29) (when $\sigma = 0$) are able to generate sustained reentrant patterns such as spiral or meander, see for example [Kee80, BKT90]. We show numerically that reentrant patterns may be generated and perpetuated only by the presence of noise. We perform the simulations on the Barkley model whose deterministic version has been intensively studied in [BKT90, Bar92, Bar94] and the model of Mitchell-Schaeffer which allows to get more realistic shape for the action potential in cardiac cells [MS03, BCF+10]. For the Barkley model, similar experiments are presented in [Sha05] where Galerkin spectral method is used as simulation scheme on a square domain. In our simulations, done on a square with periodic conditions and on a smoothed cardioid, we observe two kinds of reentrant patterns due to noise: the first may be seen as a scroll wave phenomenon whereas the second corresponds to spiral phenomenon. Both phenomena may be regarded as sources of tachyarrhythmia since in both cases, areas of the spatial domain are successively activated by the same wave which re-enters in the region.

The Barkley model reads as follows, for a, b two positive reals in $(0,1)$

$$\left\{ \begin{array}{rcl} \mathrm{d}u & = & [\nu\Delta u + \frac{1}{\varepsilon}u(1-u)(u - \frac{v+b}{a})]\mathrm{d}t + \sigma \mathrm{d}W^Q, \\ \mathrm{d}v & = & [u - v]\mathrm{d}t. \end{array} \right. \tag{1.34}$$

Let us emphasize that the model (1.34) is endowed with a timescale parameter ε. The presence of this parameter is fundamental for the observation of traveling waves in the system: ε enforces the system to be either quiescent or excited with a sharp transition between the two states. Moreover, the relative values of the timescale parameter ε and the strength of the noise σ appear to be of first importance to obtain reentrant patterns. This fact is also pointed out by our numerical bifurcation analysis. Let us mention that noise induced phenomena have been studied in [BG06] for finite dimensional systems of stochastic differential equations. The theoretical study of slow-fast SPDEs, through averaging methods, has been considered in [Bre12, CF09, WR12] for SPDEs.

All the simulations of Chapter 6 have been performed using the FreeFem++ finite element software, see [HLHOP]. This software offers the advantage to provide the mesh of the domain, the corresponding finite element basis and to solve linear problems related to the finite element discretization of the model on its own. The originality of the present work is to use this software to simulate stochastic PDEs.

1.4 Perspectives

Throughout this work, a number of areas have come to light that merit further study, they are listed below. Some questions are in progress whereas we plan to address others in future works.

Joint limit when $(N, \varepsilon) \to (\infty, 0)$. In line with the reduction of models achieved by temporal averaging in Chapters 3 and 4, one may ask the question of the reduction of the averaged model when the number of ionic channels converges to infinity. Namely, how can we treat simultaneously the phenomena of acceleration of the speed of the fast component ($\varepsilon \to 0$) and the increase of the number of ionic channels ($N \to \infty$)? For example, do we obtain the same models in permuting the order of these two different limits ?
It is worth noting that the averaging of conductance-based neuron models with multiple timescales leads to consider reaction terms for the evolution of the action potential of the form

$$G_r(u) = \frac{1}{N}\sum_{i=1}^{N} c_{\bar{r}(i)}(v_{\bar{r}(i)} - (u, \phi_{z_i}))\lambda((u, \phi_{z_i}))\phi_{z_i}.$$

These kinds of reaction terms are not considered in [Aus08] or [RTW12] where limit theorems are obtained when N goes to infinity and ε is fixed. The presence of the supplementary function λ prevents us to use the results of these papers. However, we believe that the method used in [Aus08] can be extended to this framework.

Refinements for conductance-based models. An other line of work about the reduction proposed in Chapters 3 and 4 is to consider ionic channels which possess slow and fast jump rates as we have done but this time with the inclusion of absorbing or transient states in the spirit of Sections 6.3 and 6.4 of [YZ98]. This certainly results in more tedious calculations to obtain averaging results when dealing with this kind of refinements for our model. One may also consider countably many ionic channels: in this case, the equation on the potential must be modified to incorporate this fact, one can not write $\frac{1}{N}\sum_i 1_\xi(r((i))\delta_{z_i}$ any longer but one may think of any finite measure ν on I replacing this sum of Dirac mass in writing $\int_I 1_\xi(r(x))\nu(\mathrm{d}x)$.

Fluctuations in N, speed of propagation of the action potential. Theorem 1.1.1 says that the stochastic conductance-based model and the deterministic one are consistent in the sense that when the number of ionic channels goes to infinity, the stochastic model converges towards the deterministic one. Then in [RTW12], the authors show that the variation of the empirical proportion of open ionic channels towards the deterministic probability is Gaussian when appropriately renormalized. An interesting question is to address the fluctuation of the stochastic potential u^N around its deterministic limit u. A careful reading of the proof of Theorem 1.1.1 in [Aus08] and some supplementary work show that

$$\sup_{t\in[0,T]} \mathbb{E}(\|u_t^N - u_t\|^2_{L^2(I)}) = \mathrm{O}\left(\frac{1}{N}\right).$$

This means that the limit of the process $\sqrt{N}(u^N - u)$ should exist in law when N goes to infinity. Of course, this result has to be proved rigorously by tightness arguments plus identification of the limit. Let us remark that this result has been proved recently for more regular spatially extended stochastic conductance-based models in [RT13] but not for models with Dirac masses. This result would explain the observed variation between the stochastic velocity of the nerve impulse c_N and the deterministic one c. Indeed, in the averaged Hodgkin-Huxley model without potassium dynamic for example, numerical simulations show that $c_N - c = \mathrm{O}\left(\frac{1}{\sqrt{N}}\right)$. Moreover, in this case, we observed not only that $c_N - c = \mathrm{O}\left(\frac{1}{\sqrt{N}}\right)$ but rather that $c_N - c = -\alpha\frac{1}{N}$ for some positive α. Why is the stochastic celerity smaller than the deterministic celerity? Can we mathematically prove this fact?

Further analysis for multiscale Hilbert-valued PDMPs. Theorem 1.3.3 shows that we can average stochastic conductance-based neuron models when both the potential and some ionic species evolve at fast rate. The averaged model is then simply a continuous time Markov chain. However, for this result to be completely satisfying, we have to understand better the properties of the invariant measure μ associated to the fast motion in the model. Indeed, the jump rates of the averaged

model are computed thanks to the invariant measure μ of the fast component. In a more general setting, this raises the question of properties of invariant measures for infinite dimensional PDMPs. The analogous question is investigated in [BLMZ12] for PDMPs in finite dimension: the support of the invariant measure is characterized and under Hörmander type conditions, the convergence to equilibrium is proved in total variation.
Always in the framework of neuron models with accelerated potential, one may also consider, and this is certainly the case for real neural cells, that the potential and some ionic species evolves at a faster timescale than other ionic species but with different speed of accelerations, let say ε_1 and ε_2. Is there interesting regimes for which one obtains other limits than the averaged results obtained in Chapter 5? One may think of regimes such as $\frac{\varepsilon_1}{\varepsilon_2} \to c \geq 0$ but other regimes could be addressed.

Effect of noise on excitable systems. For the investigations on the effect of noise on excitable systems for cardiac cells in Chapter 6, a lot of questions remain. For example, if we consider the Barkley model under super-threshold forcing, one may ask the influence of noise on the speed of propagation of the nucleated wave. Is the mean speed of the wave smaller than the speed of the wave without noise, as in the case of the stochastic Hodgkin-Huxley model without potassium? Is the mean speed of the wave equal to the speed of the wave in the corresponding deterministic situation?
If we consider a periodic forcing of this model, one may also ask how the periodicity of the wave train generated by the periodic forcing is affected by noise. We expect to observe, as in [TJ10] for the case with spatial dimension one, the annihilation by weak noise of the propagation of some generated waves.
In the regime where noise induced reentrant patterns have been observed, one may want to determine numerically the periodicity of the generated reentrant waves as well as the motion of the reactivation zone. This should result in a better understanding of the observed phenomena.
On a more theoretically oriented point of view, we have conducted the error analysis of the considered discretization method for stochastic linear equations in the strong sense. We believe that it should be useful to derive the strong error introduced by the scheme for the Barkley model. The literature on the error analysis in the strong sense for the finite element method for SPDEs is not very developed and such results should be appreciable. We intend to begin this analytical study on simplified excitable system such as conductance based models with only one ionic specie.

Chapter 2

Preliminary material

This chapter gathers some mathematical preliminaries used in the proof of some theorems of Chapters 3, 4 and 5. Section 2.1 is made of classical facts from functional analysis on Hilbert spaces. Then, we prove in Section 2.2 several useful estimates about the heat semigroup on a segment with zero Dirichlet boundary conditions. In Section 2.3, we recall a tightness criteria of [Mé84] for Hilbert-valued stochastic processes. We end this chapter with a reminder about Fréchet differentiability in Section 2.4 and Grönwall's lemma in Section 2.5.

2.1 An evolution triple

We set $I = [0, 1]$. Let us define:

$$\|u\|_V = \sqrt{\int_I (u(x))^2 + (u'(x))^2 \mathrm{d}x}, \quad \|u\|_{L^2(I)} = \sqrt{\int_I (u(x))^2 \mathrm{d}x}.$$

In this chapter we will work with the following triplet of Banach spaces (evolution triple) $H_0^1(I) \subset L^2(I) \subset H^{-1}(I)$. We set $V = H_0^1(I)$, it is the completion of the set of $\mathcal{C}^\infty$ functions with compact support on I with respect to the norm $\|\cdot\|_V$. $L^2(I)$ endowed with the norm $\|\cdot\|_{L^2(I)}$ is the usual space of measurable and square integrable functions with respect to the Lebesgue measure on I. $H^{-1}(I)$ is the dual space of V and will be denoted by V^*.

We recall here a few basic results about these three spaces. $L^2(I)$ and V are two separable Hilbert spaces endowed with their usual scalar products denoted respectively by $(\cdot,\cdot)_{L^2(I)}$ and $(\cdot,\cdot)$. That is

$$\forall (u, v) \in L^2(I) \times L^2(I), \quad (u, v)_{L^2(I)} = \int_I f(x)g(x)\mathrm{d}x$$

and

$$\forall (u,v) \in V \times V, \quad (u,v) = \int_I u(x)v(x)\mathrm{d}x + \int_I u'(x)v'(x)\mathrm{d}x.$$

We denote by $< \cdot,\cdot >$ the duality pairing between V and V^*. The embeddings $V \subset L^2(I) \subset V^*$ are continuous and dense. Moreover, for any $h \in L^2(I)$ and any $u \in V$: $< h, u > \equiv (h,u)_{L^2(I)}$.

For an integer $k \geq 1$ we define the following functions on I:

$$f_k(\cdot) = \sqrt{2}\sin(k\pi\cdot), \quad e_k(\cdot) = \frac{\sqrt{2}}{\sqrt{1+(k\pi)^2}}\sin(k\pi\cdot).$$

The family $\{f_k, k \geq 1\}$ (resp. $\{e_k, k \geq 1\}$) is a Hilbert basis of $L^2(I)$ (resp. V). In $L^2(I)$ (resp. V), the Laplacian with zero Dirichlet boundary conditions has the following spectral decomposition:

$$\Delta u = -\sum_{k\geq 1}(k\pi)^2(u,f_k)_{L^2(I)}f_k \quad \left(\text{resp. } \Delta u = -\sum_{k\geq 1}(k\pi)^2(u,e_k)e_k\right)$$

for u in the domain $\mathcal{D}_{L^2(I)}(\Delta) = \{u \in L^2(I); \sum_{k\geq 1} k^4(u,f_k)^2_{L^2(I)} < \infty\}$ (resp. $\mathcal{D}(\Delta) = \{u \in V; \sum_{k\geq 1} k^4(u,e_k)^2 < \infty\}$). In the two proposed basis, the above expressions mean that the laplacian operator is diagonal with eigenvalues $-(k\pi)^2$. Note that the operator Δ is symmetric on $\mathcal{D}(\Delta)$

$$\forall (u,v) \in \mathcal{D}(\Delta) \times \mathcal{D}(\Delta), \quad (\Delta u, v) = (v, \Delta u)$$

and $-\Delta$ is V-elliptic in the meaning of

$$\forall (u,v) \in \mathcal{D}(\Delta) \times \mathcal{D}(\Delta), \quad (-\Delta u, u) \geq \pi^2 \|u\|_V^2.$$

We refer the reader to Chapter 1, Section 1.3 of [Hen81] for more details. A useful result in the sequel is the continuous embedding of V in $\mathcal{C}(I)$: we denote by C_P the constant such that, for all $u \in V$ we have:

$$\sup_{x\in I}|u(x)| \leq C_P\|u\|_V.$$

2.2 About the heat semigroup on a segment with zero Dirichlet boundary conditions

Note that the chosen Hilbert basis imposes that for $u \in \mathcal{D}(\Delta)$, Δu is zero on the boundaries of I, that is in 0 and 1. It corresponds to impose zero Dirichlet boundary conditions to the laplacian operator. These boundary conditions are in fact included in the domain $\mathcal{D}(\Delta)$ of the laplacian operator. Indeed, one can show that $\mathcal{D}(\Delta) = H^2(I) \cap H_0^1(I)$. Here, $H^2(I)$ is the space of twice differentiable functions (in the distributional sens) which are in $L^2(I)$ and whose first and twice derivatives are also in $L^2(I)$. Let us define $(e^{\Delta t}, t \in \mathbb{R}_+)$, the semigroup associated to Δ in V. For any $u \in V$:

$$e^{\Delta t} u = \sum_{k \geq 1} e^{-(k\pi)^2 t}(u, e_k)e_k, \quad t > 0. \tag{2.1}$$

$(e^{\Delta t}, t \in \mathbb{R}_+)$ is a $\mathcal{C}_0$-semigroup (also referred as strongly continuous one parameter semigroup) on V with generator Δ. This means that it is a one parameter semigroup

i) $e^{\Delta 0}$ is the identity operator on V;

ii) $\forall (s,t) \in \mathbb{R}_+ \times \mathbb{R}_+, \quad e^{\Delta(t+s)} = e^{\Delta t} e^{\Delta s}$,

which is strongly continuous

iii) $\forall u \in V, \quad \lim_{t \to 0} \|e^{\Delta t} u - u\|_V = 0$.

Moreover, Δ is its infinitesimal generator

iv) $\forall u \in \mathcal{D}(\Delta), \quad \lim_{t \to 0} \frac{e^{\Delta t} u - u}{t} = \Delta u$.

As defined by (2.1), the semigroup $(e^{\Delta t}, t \in \mathbb{R}_+)$ only acts on functions of V. We define the semigroup at time t against a Dirac mass δ_y at a point y in I by

$$e^{\Delta t} \delta_y = \sum_{k \geq 1} e^{-(k\pi)^2 t}(1 + (k\pi)^2) e_k(y) e_k.$$

That is we used the representation $< \delta_y, e_k >= (1 + (k\pi)^2) e_k(y)$. We proceed to recall a lemma in [Aus08]. This lemma tells us how to estimate the integral of a function against functions of the form $e^{\Delta t} \delta_y$. Note that the last part of conclusion 1. of Lemma 2.2.1 is contained in the proof of the corresponding lemma in [Aus08]. Comparing with [Aus08], we propose a completely different proof of this lemma.

Lemma 2.2.1. *Let y be in the interior of I. Then $e^{\Delta t} \delta_y$ is a smooth function on I vanishing at the end-points for any $t > 0$. Furthermore:*

1. *There is some constant $C_1 > 0$, depending on T but otherwise not on $t \in [0, T]$, such that for any continuous function $u : [0, t] \to \mathbb{R}$ we have that the function*

$$I \to \mathbb{R} : x \mapsto \int_0^t u(s) e^{\Delta(t-s)} \delta_y(x) \mathrm{d}s$$

is in V and satisfies the estimate

$$\left\| \int_0^t u(s) e^{\Delta(t-s)} \delta_y(\cdot) \mathrm{d}s \right\|_V \leq C_1 \|u\|_\infty.$$

Moreover for any $\eta > 0$ we can choose $\varepsilon > 0$ so small that:

$$\left\| \int_{t-\varepsilon}^t u(s) e^{\Delta(t-s)} \delta_y(\cdot) \mathrm{d}s \right\|_V \leq \frac{1}{2} \eta \|u\|_\infty.$$

2. *For any fixed $\varepsilon > 0$ there is some constant $C_2(\varepsilon)$, depending on ε and T but otherwise not on $t \in [0, T]$, such that for any continuous function $u : [0, t] \to \mathbb{R}$ we have*

$$\left\| \int_0^{t-\varepsilon} u(s) e^{\Delta(t-s)} \delta_y(\cdot) \mathrm{d}s \right\|_V \leq C_2(\varepsilon) \int_0^{t-\varepsilon} |u(s)| \mathrm{d}s$$

for any $t \in [0, T]$.

Proof. Since there is no ambiguity in this proof, we simply write $\|\cdot\|$ for $\|\cdot\|_V$. Let us recall that by definition, for $t \in [0, T]$ and y in the interior of I

$$e^{\Delta t} \delta_y = \sum_{k \geq 1} e^{-(k\pi)^2 t} (1 + (k\pi)^2) e_k(y) e_k$$

with

$$e_k(y) = \sqrt{\frac{2}{1 + (k\pi)^2}} \sin(k\pi y).$$

For $\alpha, \beta \in [0, T]$, we define:

$$\begin{aligned} A(\alpha, \beta) &= \int_\alpha^\beta u(s) e^{\Delta(t-s)} \delta_y \mathrm{d}s \\ &= \sum_{k \geq 1} \sqrt{2(1 + (k\pi)^2)} \sin(k\pi y) \int_\alpha^\beta u(s) e^{-(k\pi)^2(t-s)} \mathrm{d}s e_k. \end{aligned}$$

We obtain the first inequality without difficulty, simply by estimating the term

$u(s)$ inside the integral by $\|u\|_\infty = \sup_{s\in[0,T]} |u(s)|$:

$$\begin{aligned}
\|A(0,t)\|^2 &= \left\| \sum_{k\geq 1} \sqrt{2(1+(k\pi)^2)} \sin(k\pi y) \int_0^t u(s) e^{-(k\pi)^2(t-s)} \mathrm{d}s e_k \right\|^2 \\
&= 2\sum_{k\geq 1} (1+(k\pi)^2) \sin^2(k\pi y) \left(\int_0^t u(s) e^{-(k\pi)^2(t-s)} \mathrm{d}s \right)^2 \\
&\leq 2\sum_{k\geq 1} (1+(k\pi)^2) \sin^2(k\pi y) \|u\|_\infty^2 \left(\frac{1-e^{-(k\pi)^2 t}}{(k\pi)^2} \right)^2 \\
&\leq 2\|u\|_\infty^2 \sum_{k\geq 1} \frac{(1+(k\pi)^2)}{(k\pi)^4} \sin^2(k\pi y).
\end{aligned}$$

We choose C_1 such that:

$$C_1^2 = 2\sum_{k\geq 1} \frac{(1+(k\pi)^2)}{(k\pi)^4} \sin^2(k\pi y),$$

to obtain the first inequality

$$\|A(0,t)\| \leq C_1 \|u\|_\infty.$$

Let us consider $\varepsilon > 0$ and write

$$\begin{aligned}
\|A(t-\varepsilon,t)\|^2 &= \left\| \sum_{k\geq 1} \sqrt{2(1+(k\pi)^2)} \sin(k\pi y) \int_{t-\varepsilon}^t u(s) e^{-(k\pi)^2(t-s)} \mathrm{d}s e_k \right\|^2 \\
&= 2\sum_{k\geq 1} (1+(k\pi)^2) \sin^2(k\pi y) \left(\int_{t-\varepsilon}^t u(s) e^{-(k\pi)^2(t-s)} \mathrm{d}s \right)^2 \\
&\leq 2\sum_{k\geq 1} (1+(k\pi)^2) \sin^2(k\pi y) \|u\|_\infty^2 \left(\frac{1-e^{-(k\pi)^2 \varepsilon}}{(k\pi)^2} \right)^2 \\
&\leq 2\|u\|_\infty^2 \sum_{k\geq 1} \frac{(1+(k\pi)^2)}{(k\pi)^4} \sin^2(k\pi y) \left(1-e^{-(k\pi)^2\varepsilon}\right)^2.
\end{aligned}$$

Let us define

$$C_1(\varepsilon) = 2\sum_{k\geq 1} \frac{(1+(k\pi)^2)}{(k\pi)^4} \sin^2(k\pi y) \left(1-e^{-(k\pi)^2\varepsilon}\right)^2.$$

By dominated convergence, for any $\eta > 0$ we can choose ε so small that $C_1(\varepsilon) \leq \left(\frac{1}{2}\eta\right)^2$. Hence:

$$\|A(t-\varepsilon, t)\| \leq \frac{1}{2}\eta\|u\|_\infty.$$

We are left with the last inequality, let $\varepsilon \in]0, t]$ and $t \in [0, T]$:

$$\begin{aligned}
\|A(0, t-\varepsilon)\|^2 &= \left\| \sum_{k\geq 1} \sqrt{2(1+(k\pi)^2)} \sin(k\pi y) \int_0^{t-\varepsilon} u(s) e^{-(k\pi)^2(t-s)} \mathrm{d}s e_k \right\|^2 \\
&= 2\sum_{k\geq 1}(1+(k\pi)^2)\sin^2(k\pi y)\left(\int_0^{t-\varepsilon} u(s)e^{-(k\pi)^2(t-s)}\mathrm{d}s\right) \\
&\leq 2\sum_{k\geq 1}(1+(k\pi)^2)\sin^2(k\pi y)e^{-(k\pi)^2\varepsilon}\left(\int_0^{t-\varepsilon} u(s)ds\right)^2.
\end{aligned}$$

The last inequality is obtained by setting

$$C_2^2(\varepsilon) = 2\sum_{k\geq 1}(1+(k\pi)^2)\sin^2(k\pi y)e^{-(k\pi)^2\varepsilon}.$$

□

2.3 Tightness in Hilbert spaces

In this part we are interested in a criteria of tightness for a family of Hilbert space valued processes. Indeed, in the sequel, we will want to prove the convergence in law of some family of processes $(u^\varepsilon, \varepsilon \in]0, 1])$ valued in $\mathcal{C}([0, T], H)$ when ε goes to zero for some Hilbert space H. We begin by recalling a criteria of relative compactness in Hilbert space. We provide a proof since we did not find any adapted references in the literature for this theorem. Moreover, this is a pleasant exercise.

Theorem 2.3.1. *Let H be a separable Hilbert space and $\{e_k, k \geq 1\}$ a Hilbert basis of H. A subset C of H has compact closure in H if and only if*

i) C is bounded in H.

ii) For any $\eta > 0$, there exists a N such that

$$\sup_{x\in H} \|x - \Pi_N x\| < \eta.$$

Π_N is the projection of x on the space $H_N = \mathrm{span}\{e_k, 1 \leq k \leq N\}$.

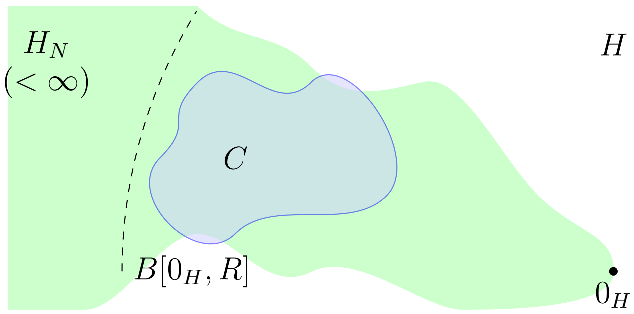


Figure 2.1: Relative compactness in Hilbert spaces.

Proof. Let us begin with the "if" part. Let $\eta > 0$ and C be a of compact closure in H. We consider the following covering of C by open balls of radius $\frac{1}{3}\eta$

$$C \subset \bigcup_{h \in C} \mathrm{B}\left(h, \frac{1}{3}\eta\right).$$

Since C is of compact closure in H, there exists $h_1, \dots, h_n$, n elements of C such that

$$C \subset \bigcup_{k=1}^{n} \mathrm{B}\left(h_k, \frac{1}{3}\eta\right).$$

Then, for any $k \in \{1, \dots, n\}$, there exists $N(k) \in \mathbb{N}$ such that

$$\|h_k - \Pi_{N(k)} h_k\| < \frac{1}{3}\eta.$$

Let $N^* = \max_{k \in \{1,\dots,n\}} N(k)$. For any $h \in C$, there is a $\tilde{k} \in \{1, \dots, n\}$ such that

$$h \in \mathrm{B}\left(h_{\tilde{k}}, \frac{1}{3}\eta\right).$$

Thus

$$\begin{aligned}
\|h - \Pi_{N^*} h\| &\leq \|h - h_{\tilde{k}}\| + \|h_{\tilde{k}} - \Pi_{N^*} h_{\tilde{k}}\| + \|\Pi_{N^*} h_{\tilde{k}} - \Pi_{N^*} h\| \\
&\leq 2\|h - h_{\tilde{k}}\| + \|h_{\tilde{k}} - \Pi_{N^*} h_{\tilde{k}}\| \\
&\leq 2\frac{\eta}{3} + \frac{\eta}{3} \\
&= \eta.
\end{aligned}$$

We go on with the "only if" part. The aim of the proof is to show that when conditions i) and ii) are fulfilled the set C is of compact closure. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of C. We want to show that this sequence has a convergent subsequence. Let $\eta > 0$, by assumption ii), there exists an integer N such that

$$\sup_{x\in C} \|x - \Pi_N x\| \leq \frac{1}{3}\eta.$$

By assumption i), C is bounded, thus, for any $k \geq 1$, the sequence $((x_n, e_k))_{n\in\mathbb{N}}$ is bounded in $\mathbb{R}$ since there is a constant α (independent of n) such that

$$\sum_{k\geq 1} (x_n, e_k)^2 \leq \alpha.$$

Therefore, the sequence $((x_n, e_k))_{n\in\mathbb{N}}$ has a Cauchy sub-sequence. Let us denote by $((x_{n_p}, e_k))_{p\in\mathbb{N}}$ this sub-sequence. There is an integer P_k such that for any $p, q \geq P_k$

$$|(x_{n_p}, e_k) - (x_{n_q}, e_k)| \leq \frac{1}{2\alpha N}\frac{1}{9}\eta^2.$$

Let us set $P^* = \max_{k\in\{1,\dots,N\}} P_k$. For any $p, q \geq P^*$ we have

$$\|x_{n_p} - x_{n_q}\| \leq \|x_{n_p} - \Pi_N x_{n_p}\| + \|\Pi_N x_{n_q} - x_{n_q}\| + \|\Pi_N x_{n_p} - \Pi_N x_{n_q}\|.$$

Then for $r = p, q$

$$\|x_{n_r} - \Pi_N x_{n_r}\| \leq \frac{1}{3}\eta$$

and

$$\|\Pi_N x_{n_p} - \Pi_N x_{n_q}\|^2 = \sum_{k=1}^{N} ((x_{n_p}, e_k) - (x_{n_q}, e_k))^2 \leq 2\alpha N \frac{1}{2\alpha N}\frac{1}{9}\eta^2 = \frac{1}{9}\eta^2.$$

Assembling the above inequalities yields

$$\|x_{n_p} - x_{n_q}\| \leq \eta.$$

This means that the sequence $(x_{n_p})_{p\in\mathbb{N}}$ is a Cauchy sequence in H which is a complete metric space. The result follows. □

From the above characterization of compactness in a Hilbert space, we derive the following characterization of tightness of a family of Hilbert space valued random variables. The two theorems below come from [Mé84].

Theorem 2.3.2 (Tightness in a Hilbert space)**.** *Let H be a separable Hilbert space endowed with the basis $\{e_k, k \geq 1\}$. We denote by H_N, for $N \geq 1$*

$$H_N = \mathrm{span}\{e_k, 1 \leq k \leq N\}.$$

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space satisfying the usual conditions and $\{u^\varepsilon, \varepsilon \in]0,1]\}$ a family of H valued random variables. Then the family $\{u^\varepsilon, \varepsilon \in]0,1]\}$ is tight in H if, and only if, for any $\delta, \eta > 0$ there is $\rho, \varepsilon_0 > 0$ and a subspace $H_{\delta,\eta}$ of $\{H_N, N \geq 1\}$ such that

1.

$$\sup_{\varepsilon\in]0,\varepsilon_0]} \mathbb{P}(\|u^\varepsilon\|_H > \rho) \leq \delta, \tag{2.2}$$

2.

$$\sup_{\varepsilon\in]0,\varepsilon_0]} \mathbb{P}(d(u^\varepsilon, H_{\delta,\eta}) > \eta) \leq \delta, \tag{2.3}$$

where $d(u_t^\varepsilon, H_{\delta,\eta}) = \inf_{v\in,H_{\delta,\eta}} \|u^\varepsilon - v\|_H$.

Proof. By Theorem 2.3.1, if the family $\{u^\varepsilon, \varepsilon \in]0,1]\}$ is tight in H then the two conditions (2.2) and (2.3) are obviously fulfilled. To prove the converse, we assume that the conditions (2.2) and (2.3) are fulfilled. Let $\delta, \varepsilon_0 > 0$ and consider the set

$$C_\delta = H \setminus \left(\bigcup_{k\in\mathbb{N}} \{g \in H;\ \|g\| > \rho_\delta\} \cup \{g \in H;\ d(g, H_{\frac{\delta}{2^k},\frac{1}{k}}) > \frac{1}{k}\} \right).$$

Then, by Theorem 2.3.1, C_δ is of compact closure in H and for any $\varepsilon \in]0, \varepsilon_0]$

$$\mathbb{P}(u^\varepsilon \in C_\delta) \geq 1 - 2\delta,$$

which means that the family $\{u^\varepsilon, \varepsilon \in]0,1]\}$ is tight in H. □

Let T be a finite time horizon. For any separable Hilbert space H (or more generally any complete and separable space), the space $\mathbb{D}([0,T], H)$ of càdlàg paths from $[0,T]$ into H is complete and separable when endowed with its Skorohod topology. This space of paths is treated comprehensively in Chapter 3 of [EK86]. The following theorem recalls that tightness of a family of processes at each fixed time t plus Aldous's condition imply the tightness of the family [Ald78].

Theorem 2.3.3 (General criterion for tightness)**.** *On the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$, *we consider* $\{u^\varepsilon, \varepsilon \in]0,1]\}$ *a family of* $\mathbb{D}([0,T], H)$ *valued random variables for* H *a separable Hilbert space. Let us assume that* $\{u^\varepsilon, \varepsilon \in]0,1]\}$ *satisfies Aldous's condition which means that for any* $\delta, M > 0$, *there exist* $\eta, \varepsilon_0 > 0$ *such that for all stopping times* τ *such that* $\tau + \eta < T$*:*

$$\sup_{\varepsilon\in]0,\varepsilon_0]} \sup_{\theta\in]0,\eta[} \mathbb{P}(\|u^\varepsilon_{\tau+\theta} - u^\varepsilon_\tau\|_H \geq M) \leq \delta. \tag{2.4}$$

If moreover, for each $t \in [0,T]$ *fixed the family* $\{u_t^\varepsilon, \varepsilon \in]0,1]\}$ *is tight in* H *then* $(u^\varepsilon, \varepsilon \in]0,1])$ *is tight in* $\mathbb{D}([0,T], H)$.

2.4 Fréchet differentiability

In this section, H is a separable Hilbert space. We say that a function $f : H \to \mathbb{R}$ has a Fréchet derivative in $u \in H$ if there exists a bounded linear operator $T_u : H \mapsto \mathbb{R}$ such that

$$\lim_{h\to 0} \frac{f(u+h) - f(u) - T_u(h)}{\|h\|_H} = 0.$$

We then write $\frac{\mathrm{d}f}{\mathrm{d}u}(u)$ for the operator T_u. For example, with $H = L^2(I)$, the square of the $\|\cdot\|_{L^2(I)}$-norm and the Dirac distribution in $x \in I$ are Fréchet differentiable on $L^2(I)$. For all $u \in L^2(I)$,

$$\frac{\mathrm{d}\|\cdot\|^2_{L^2(I)}}{\mathrm{d}u}(u)[h] = 2(u,h), \quad \frac{\mathrm{d}\delta_x}{\mathrm{d}u}(u)[h] = h(x)$$

for all $h \in L^2(I)$. Let us consider $\{e_k, k \geq 1\}$ a Hilbert basis of H and $\{e_k^*, k \geq 1\}$ the corresponding dual basis: $e_k^*(e_p) = \delta_{kp}$ (Kronecker symbol) for $(k,p) \in \mathbb{N}\times\mathbb{N}$. Then, we see without difficulty that for any $u \in H$

$$\frac{\mathrm{d}f}{\mathrm{d}u}(u) = \sum_{k\geq 1} \partial_k f(u) e_k^*,$$

where by definition $\partial_k f(u) = \frac{\mathrm{d}f}{\mathrm{d}u}(u)[e_k]$ for $k \geq 1$. In the same way, we can define the Fréchet derivative of order 2. The second Fréchet derivative of a twice Fréchet differentiable function $f : H \to \mathbb{R}$ is denoted by $\frac{\mathrm{d}^2 f}{\mathrm{d}u^2}(u)$. It can be considered as a bilinear form on $H \times H$. For instance, with $H = V$,

$$\frac{\mathrm{d}^2\|\cdot\|^2_V}{\mathrm{d}u^2}(u)[h,k] = 2(h,k), \quad \frac{\mathrm{d}^2\delta_x}{\mathrm{d}u^2}(u)[h,k] = 0$$

for all $(h,k) \in V \times V$. Note that $\{e_k \otimes e_p, k, p \geq 1\}$ is a Hilbert basis of $H \otimes H$ where we recall that $(u \otimes v, \tilde{u} \otimes \tilde{v}) = (u, \tilde{u})(v, \tilde{v})$. The dual basis $\{e_k^* \otimes e_p^*, k, p \geq 1\}$ is such that $(e_k^* \otimes e_p^*, e_i \otimes e_j) = \delta_{ki}\delta_{pj}$. Then

$$\frac{\mathrm{d}^2 f}{\mathrm{d}u^2}(u) = \sum_{k,p\geq 1} \partial_{kp} f(u) e_k^* \otimes e_p^*,$$

where by definition $\partial_{kp} f(u) = \frac{\mathrm{d}^2 f}{\mathrm{d}u^2}(u)[e_k, e_p]$ for $k, p \geq 1$. Fréchet differentiation is stable by sum and product.

2.5 Grönwall's lemma

At numerous points in the text, we will be lead to bound the current value of some deviation between two processes in terms of some average of the values this deviation has taken so far. The most common formalization of this idea is Grönwall's lemma.

Proposition 2.5.1. *Suppose $T > 0$ and $f : [0, T] \to \mathbb{R}$ is continuous. Suppose further that there are two positive constants A, B such that*

$$f(t) \leq A + B \int_0^t f(s)\mathrm{d}s$$

for all $t \in [0, T]$. Then $f(t) \leq Ae^{Bt}$ for all $t \in [0, T]$.

In Chapter 6, we will use the same idea in the discrete setting. The following result is sometimes referred to as discrete Grönwall's lemma.

Proposition 2.5.2. *Suppose $N \in \mathbb{N}$ and $(x_n)_{n\geq 0}$ and $(y_n)_{n\geq 0}$ are positive real sequences. Suppose further that there are two positive constants a, b such that*

$$x_{n+1} \leq ax_n + by_n$$

for all $n \in \{0, \ldots, N-1\}$. Then

$$x_n \leq a^n x_0 + b \sum_{k=0}^{n-1} a^{n-k-1} y_k$$

for all $n \in \{0, \ldots, N\}$.

Chapter 3

Averaging for a fully coupled piecewise-deterministic Markov process in infinite dimensions

The material for chapter 3 is taken from the article [GT12] *Averaging for a Fully-Coupled Piecewise Deterministic Markov Process in Infinite Dimensions* published in *Advances in Applied Probability*, Volume 44, Number 3 (2012), 749-773. Since the present thesis is provided with a general introduction in Chapter 1 and especially for this chapter in Section 1.3.2, we start with a shorter introduction than in [GT12].

3.1 Introduction

The Hodgkin-Huxley model is one of the most studied models in neuroscience since its creation in the early 50's by Hodgkin and Huxley [HH52]. This deterministic model was first created in order to describe the propagation of an action potential or nerve impulse along the axon of a neuron of the giant squid. In this chapter, we will consider a mathematical generalization of the Hodgkin-Huxley model investigated numerically in [FWL05] and analytically in [Aus08]. This model, called the spatial stochastic Hodgkin-Huxley model in the sequel, is made of a partial differential equation (PDE) describing the evolution of the variation of the potential across the membrane, coupled with a continuous time Markov chain which describes the dynamics of ion channels which are present all along the axon. It is a Piecewise Deterministic Markov Process (PDMP) in infinite dimensions (see [BR11, Rie12b, RTW12] for PDMP in infinite dimensions and also [CD10, CD11, CDR09, BLBMZ12] and references therein for PDMP in finite dimension). Moreover it is fully-coupled in the sense that the evolution of the potential depends on the kinetics of ion channels and vice-versa. The role of ion channels is fundamental because they allow and amplify the propagation of an action potential. We introduce different time scales in this model considering that some ion channels open and close at faster rates than others. We perform a slow-fast analysis of this model and prove that asymptotically it reduces to the so-called averaged model which is still a PDMP in infinite dimensions for which we provide effective evolution equations and jump rates. We thus reduce the complexity of the original model by simplifying the kinetics of ion channels. To the best of our knowledge no averaging results are available for PDMP in infinite dimensions.

We conclude this short introduction with the plan of the chapter. In Section 3.2 we introduce the spatial stochastic Hodgkin-Huxley model and the formalism of PDMP in infinite dimensions. Then we introduce the two time scales in the model and prove a crucial result for the sequel. We finish this section by presenting the main assumptions on the model and our main result. Section 3.3 is devoted to the proof of our main result. In Section 3.4 we apply our result to a Hodgkin-Huxley type model as an example. At the end of the chapter, an appendix provides the parameter values used in the simulation presented in Section 3.4.

3.2 Statement of the model and results

3.2.1 The spatial stochastic Hodgkin-Huxley model

We recall here the stochastic Hodgkin-Huxley equations considered in [Aus08]. Basically this spatial stochastic Hodgkin-Huxley model describes the propagation of an action potential along an axon. The axon is the part of a neuron which transmits the information received from the soma to another neuron on long distances: the length of the axon is big relative to the size of the soma. Along the axon are the ion channels which amplify and allow the propagation of the received impulse. For mathematical convenience, we assume that the axon is a segment of length one and we denote it by $I = [0, 1]$. The ion channels are assumed to be in number $N \geq 1$ and regularly placed along the axon at the place $\frac{i}{N}$ for $i \in \mathcal{N}$. This distribution of the channels is certainly unrealistic but we assume it to fix the ideas and in accordance with [Aus08]. However the mathematics are the same if we consider any finite subset of I instead of $\frac{1}{N}\mathcal{N}$. Each ion channel can be in a state $\xi \in E$ where E is a finite states space. For example, for the Hodgkin-Huxley model, a state can be : "receptive to sodium ions and open" (see [Hil84]).

The ion channels switch between states according to a continuous time Markov chain whose jump intensities depend on the local potential of the membrane, that is why the model is said to be fully-coupled. For any states $\xi, \zeta \in E$ we define by $\alpha_{\xi,\zeta}$ the jump intensity (or rate function) from the state ξ to the state ζ. It is a real valued function of a real variable and supposed to be, as is its derivative, Lipschitz-continuous. We assume moreover that: $0 \leq \alpha_{\xi,\zeta} \leq \alpha^+$ for any $\xi, \zeta \in E$ and either $\alpha_{\xi,\zeta}$ is constant equal to zero or is strictly positive bounded below by a strictly positive constant α_-. That is, the non-zero rate functions are bounded below and above by strictly positive constants. For a given channel, the rate function describes the rate at which it switches from one state to another.

A possible configuration of all the N ion channels is denoted by an element $r = (r(i), i \in \mathcal{N})$ of $\mathcal{R} = E^{\mathcal{N}}$: $r(i)$ is the state of the channel which is at the position $\frac{i}{N}$ for $i \in \mathcal{N}$. The channels, or stochastic processes $r(i)$, are supposed to evolve independently over infinitesimal timescales. Denoting by $u_t(\frac{i}{N})$ the local potential at point $\frac{i}{N}$ at time t, we have:

$$\mathbb{P}(r_{t+h}(i) = \zeta | r_t(i) = \xi) = \alpha_{\xi,\zeta}\left(u_t\left(\frac{i}{N}\right)\right)h + \mathrm{o}(h). \tag{3.1}$$

For any $\xi \in E$ we also define the maximal conductance c_ξ and the steady state potentials v_ξ of a channel in state ξ which are both constants, the first being positive.

The membrane potential $u_t(x)$ along the axon evolves according to the following PDE:

$$\partial_t u_t = \Delta u_t + \frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}} c_\xi 1_\xi(r_t(i))(v_\xi - u_t(\frac{i}{N}))\delta_{\frac{i}{N}}. \tag{3.2}$$

We will assume zero Dirichlet boundary conditions for this PDE (clamped axon). We are interested in the process $(u_t, r_t)_{t\in[0,T]}$.

We recall here the result of [Aus08] which states that there exists a stochastic process satisfying equations (3.2) and (3.1). Let u_0 be in $\mathcal{D}(\Delta)$ such that $\min_{\xi\in E} v_\xi \leq u_0 \leq \max_{\xi\in E} v_\xi$, the initial potential of the axon. Let $q_0 \in \mathcal{R}$ be the initial configuration of ion channels. Let us also recall that, in accordance with the notations of Chapter 2, $V = H_0^1(I)$ in the present chapter.

Proposition 3.2.1 ([Aus08]). *Fix $N \geq 1$ and let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions. There exist a unique pair $(u_t, r_t)_{0\leq t\leq T}$ of càdlàg adapted stochastic processes such that each sample path of u is a continuous map from $[0,T]$ to V and r_t is in $\mathcal{R}$ for all $t \in [0,T]$ and satisfying:*

- *(Regularity) The map $t \mapsto \partial_t u_t$ lies in $L^2([0,T], V^*)$ $\mathbb{P}$-almost surely.*
- *(Dynamic: PDE)*

$$\partial_t u_t = \Delta u_t + \frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}} c_\xi 1_\xi(r_t(i))(v_\xi - u_t(\frac{i}{N}))\delta_{\frac{i}{N}}$$

$\forall t \in [0,T]$, $\mathbb{P}$-*almost surely.*

- *(Dynamic: jump)*

$$\mathbb{P}(r_{t+h}(i) = \zeta | r_t(i) = \xi) = \alpha_{\xi,\zeta}\left(u_t\left(\frac{i}{N}\right)\right)h + \mathrm{o}(h).$$

- *(Initial condition: PDE) u_0 given.*
- *(Initial condition: jump) q_0 given.*
- *(Boundary conditions: PDE only) $u_t(0) = u_t(1) = 0$, $\forall t \in [0, T]$.*

Moreover there exists a constant C such that $\|u_t\|_\infty \leq C$ for all $t \in [0, T]$ and ω.

In [Aus08] the author proves that when the number of ion channels increases to infinity, the above model converges, in a sense, towards a deterministic model. In our paper, unlike in [Aus08] we will work with a fixed number of ion channels but introduce two timescales in the model in order to perform a slow-fast analysis.

3.2.2 The spatial stochastic Hodgkin-Huxley model as a Piecewise Deterministic Markov Process in infinite dimensions

The paper [BR11] extends the theory of finite dimensional Piecewise Deterministic Markov Processes (PDMP) introduced in [Dav84] and [Dav93] to PDMP in infinite dimensions. The results stated here come from section 2.3 and 3.1 of [BR11] adapted to our particular notations.

For all $r \in \mathcal{R}$ we define the following function on V:

$$G_r(u) = \frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}} c_\xi 1_\xi(r(i))(v_\xi - u(\frac{i}{N}))\delta_{\frac{i}{N}}, \tag{3.3}$$

which is V^* valued. We do not write the dependence in N in the notation of $G_r(u)$ since, unlike in [Aus08], N is a fixed parameter here.

The stochastic process $(u_t, r_t)_{t\in[0,T]}$ can be described in the following way. We start with a given initial potential v_0 and ion channels configuration q_0. Then the u component evolves according to the evolution equation:

$$\partial_t u_t = \Delta u_t + G_{q_0}(u_t),$$

until the first jump of the r component at time τ_1 which occurs according to the transition rate function $q : V \times \mathcal{R} \to \mathbb{R}_+$ given for $(u, r) \in V \times \mathcal{R}$ by:

$$q_r(u) = \sum_{i\in\mathcal{N}}\sum_{\xi\neq r(i)} \alpha_{r(i),\xi}(u(\frac{i}{N})),$$

according to (3.1). A new value q_1 for r is then chosen according to a transition measure J from $V \times \mathcal{R}$ to $\mathcal{P}(\mathcal{R})$ (the set of all probability measures on $\mathcal{R}$). This transition measure is also given by the jump distribution of the ion channels. For $(u, r) \in V \times \mathcal{R}$ and $r' \in \mathcal{R}$ which differs from r only by the component $r(i_0)$:

$$J_{(u,r)}(\{r'\}) = \frac{\alpha_{r(i_0),r'(i_0)}(u(\frac{i_0}{N}))}{\Lambda(u, r)}.$$

If r' differs from r by two or more components then $J_{(u,r)}(\{r'\}) = 0$.

Then u evolves according to the "updated" evolution equation:

$$\partial_t u_t = \Delta u_t + G_{q_1}(u_t)$$

with initial condition u_{τ_1} and Dirichlet boundary conditions, until the second jump of the r component. And so on. This description justifies the term "Piecewise Deterministic Process": the dynamic of the process is indeed purely deterministic between the jumps.

As shown in [BR11] in Theorem 6, the equations of Proposition 3.2.1 define a standard Piecewise Deterministic Process (PDP) in the sense of Definition 3 of [BR11]. We recall briefly here what this means technically:

- for a given ion channel configuration $r \in \mathcal{R}$, the PDE:

  $$\partial_t u_t = \Delta u_t + G_r(u_t)$$

  with zero Dirichlet boundary conditions admits a unique global (weak) solution continuous on V for every initial value $u_0 \in V$ denoted by $\psi_r(t, x)$.

- The number of state switches in the $(r_t)_{t \in [0,T]}$ component during any finite time interval is finite almost surely.

- The transition rate function $q : (u, r) \in V \times \mathcal{R} \mapsto q_r(u) \in \mathbb{R}_+$ from one state of $\mathcal{R}$ to another is a measurable function and for all $(u, r) \in V \times \mathcal{R}$, the transition rate as a function of time is integrable over every finite time interval but diverges to ∞ when it is integrated over $\mathbb{R}_+$ (the expected number of jumps tends to ∞ when the time horizon increases).

- The transition measure J from one state of $\mathcal{R}$ to another is such that the mapping $(u, r) \mapsto J_{(u,r)}(R)$ is measurable for every $R \subset \mathcal{R}$ and

  $$J_{(u,r)}(\{r\}) = 0$$

  for all $(u, r) \in V \times \mathcal{R}$.

In fact the stochastic process $(u_t, r_t)_{t\in[0,T]}$ is more than a piecewise deterministic process, it is a piecewise deterministic *Markov* process.

Theorem 3.2.1. *1. Our infinite-dimensional standard PDP is a homogeneous Markov process on $V \times \mathcal{R}$.*

2. Locally bounded measurable functions on $V \times \mathcal{R}$ which are absolutely continuous as maps $t \mapsto f(\psi_r(t,x), r)$ for all (x,r) are in the domain $\mathcal{D}(\mathcal{A})$ of the extended generator. The extended generator is given for almost all t by:

$$\mathcal{A}f(u_t, r_t) = \frac{\mathrm{d}}{\mathrm{d}t} f(u_t, r_t) + \mathcal{B}f(u_t, r_t), \tag{3.4}$$

where

$$\mathcal{B}f(u_t, r_t) = \sum_{i\in\mathcal{N}} \sum_{\zeta\in E} [f(u_t, r_t(r_t(i) \to \zeta)) - f(u_t, r_t)]\alpha_{r_t(i),\zeta}(u_t(\frac{i}{N}))$$

with $r_t(r_t(i) \to \zeta)$ is the component of $\mathcal{R}$ with $r_t(r_t(i) \to \zeta)(j)$ equals to $r_t(j)$ if $j \neq i$ and to ζ if $j = i$. The notation $\frac{\mathrm{d}}{\mathrm{d}t} f(u_t, r_t)$ means that the function $s \mapsto f(u'_s, r)$ is differentiated at $s = t$, where u' is the solution of the PDE of Proposition 3.2.1 such that $u'_t = u_t$ and with the channel state r held fixed equal to r_t.

Remark 3.2.1. *It is not usual to write a generator only along paths $(u_t, r_t)_{t\geq 0}$ as we do. We would expect for a generator an analytical expression of the form $\mathcal{A}f(u,r)$ for any $(u,r) \in V \times \mathcal{R}$. It is in fact possible to obtain such an expression for more regular function f. The part of the generator related to the continuous Markov chain r takes the form $\mathcal{B}f(u,r)$ with no changes but the derivative $\frac{\mathrm{d}}{\mathrm{d}t} f(u,r)$ can then be expressed thanks to the Fréchet derivative of f, see [BR11] for more details. We will not use this refinement in the sequel.*

Knowing that a stochastic process is Markovian is crucial for its study since a lot of mathematical tools have been developed for Markovian processes, see for example [EK86].

3.2.3 Singularly perturbed model and main results

Introduction of two time scales in the model.

We introduce now two time scales in the model. Indeed, we consider that the Hodgkin-Huxley model is a slow-fast system: some states of the ion channels communicate faster between them than others, see for example [Hil84]. Mathematically, this leads to introduce a small parameter $\varepsilon > 0$ in our equation. For the states which communicate at a faster rate, we say that they communicate at the usual rate divided by ε.

We can then consider different classes of states or partition the state space E in states which communicate at a high rate. This kind of description is very classical, see for example [FGC08]. We regroup our states in classes making a partition of the state space E into:

$$E = E_1 \sqcup \cdots \sqcup E_l,$$

where $l \in \{1, 2, \cdots\}$ is the number of classes. Inside a class E_j, the states communicate faster at jump rates of order $\frac{1}{\varepsilon}$. States of different classes communicate at the usual rate of order 1. For $\varepsilon > 0$ fixed, we denote by $(u^\varepsilon, r^\varepsilon)$ the modification of the PDMP introduced in the previous section with now two time scales. Its generator is, for $f \in \mathcal{D}(\mathcal{A})$:

$$\mathcal{A} f(u_t^\varepsilon, r_t^\varepsilon) = \frac{\mathrm{d}}{\mathrm{d}t} f(u_t^\varepsilon, r_t^\varepsilon) + \mathcal{B}^\varepsilon f(u_t^\varepsilon, r_t^\varepsilon) \tag{3.5}$$

$\mathcal{B}^\varepsilon$ is the component of the generator related to the continuous time Markov chain r^ε. According to our slow-fast description we have the two time scales decomposition of this generator:

$$\mathcal{B}^\varepsilon = \frac{1}{\varepsilon}\mathcal{B} + \hat{\mathcal{B}},$$

where the "fast" generator $\mathcal{B}$ is given by:

$$\mathcal{B} f(u_t^\varepsilon, r_t^\varepsilon) = \sum_{i \in \mathcal{N}} \sum_{j=1}^{l} 1_{E_j}(r_t^\varepsilon(i)) \sum_{\zeta \in E_j} [f(u_t^\varepsilon, r_t^\varepsilon(r_t^\varepsilon(i) \to \zeta) - f(u_t^\varepsilon, r_t^\varepsilon)] \alpha_{r_t^\varepsilon(i),\zeta}(u_t^\varepsilon(\frac{i}{N}))$$

and the "slow" generator is given by:

$$\hat{\mathcal{B}} f(u_t^\varepsilon, r_t^\varepsilon) = \sum_{i \in \mathcal{N}} \sum_{j=1}^{l} 1_{E_j}(r_t^\varepsilon(i)) \sum_{\zeta \notin E_j} [f(u_t^\varepsilon, r_t^\varepsilon(r_t^\varepsilon(i) \to \zeta) - f(u_t^\varepsilon, r_t^\varepsilon)] \alpha_{r_t^\varepsilon(i),\zeta}(u_t^\varepsilon(\frac{i}{N})).$$

For $y \in \mathbb{R}$ fixed and $g : \mathbb{R} \times E \to \mathbb{R}$, we denote by $\mathcal{B}_j(y)$, $j \in \{1, \cdots, l\}$ the following generator:

$$\mathcal{B}_j(y) g(\xi) = 1_{E_j}(\xi) \sum_{\zeta \in E_j} [g(y, \zeta) - g(y, \xi)] \alpha_{\xi,\zeta}(y),$$

which will be of interest in the sequel.

Uniform boundedness.

Here is a crucial result for the proof of our main result. Its proof takes back the argument developed in [Aus08] but for the sake of completeness and given the intensive use of this proposition in the sequel, we provide a short proof.

Proposition 3.2.2. *For any $T > 0$, there is a deterministic constant $C > 0$ independent of $\varepsilon \in]0,1]$ such that:*

$$\sup_{t\in[0,T]} \|u_t^\varepsilon\|_V \leq C,$$

almost surely.

In the proof of this proposition, we will use the following representation of a solution of the PDE part of our PDMP. We say that u^ε is a mild solution to (3.2) if:

$$u_t^\varepsilon = e^{\Delta t}u_0 + \int_0^t e^{\Delta(t-s)} G_{r_s^\varepsilon}(u_s^\varepsilon)\mathrm{d}s, \quad t \geq 0, \tag{3.6}$$

almost surely where we recall that $(e^{\Delta t}, t \geq 0)$ is the semi-group associated to Δ in V. For any $u \in V$:

$$e^{\Delta t}u = \sum_{k\geq 1} e^{-(k\pi)^2 t}(u, e_k)e_k, \quad t > 0.$$

Let us mention that the mild formulation (3.6) holds also up to a bounded stopping time τ, this will be useful later in the text.

Proof. We work ω by ω. Using the mild formulation (3.6), we see that:

$$\begin{aligned} \|u_t^\varepsilon\|_V &\leq \|e^{\Delta t}u_0\|_V + \frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}} \|\int_0^t c_\xi 1_\xi(r_s^\varepsilon(i))v_\xi e^{\Delta(t-s)}\delta_{\frac{i}{N}}\mathrm{d}s\|_V \\ &+ \frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}} \|\int_0^t c_\xi u_s^\varepsilon(\frac{i}{N}) e^{\Delta(t-s)}\delta_{\frac{i}{N}}\mathrm{d}s\|_V. \end{aligned}$$

By Lemma 2.2.1, there exists a constant C_1 depending only of T such that

$$\|\int_0^t c_\xi 1_\xi(r_s^\varepsilon(i))v_\xi e^{\Delta(t-s)}\delta_{\frac{i}{N}}\mathrm{d}s\|_V \leq \max_{\xi\in E} c_\xi|v_\xi|C_1.$$

Therefore, the two first term of the above sums are bounded by a fixed constant α_1 which can be chosen independent of $t \in [0,T]$. For the second sum, we break the integral in two pieces. For all $\gamma > 0$ by Lemma 2.2.1, we can choose $\eta > 0$ so small that:

$$\| \int_{t-\eta}^{t} c_\xi u_s^\varepsilon(\frac{i}{N}) e^{\Delta(t-s)} \delta_{\frac{i}{N}} \mathrm{d}s \|_V \leq \frac{1}{2}\gamma \sup_{s\in[0,t]} |c_\xi u_s^\varepsilon(\frac{i}{N})|.$$

Since

$$|c_\xi u_s^\varepsilon(\frac{i}{N})| \leq \max c_\xi \|u_s^\varepsilon\|_\infty \leq \max c_\xi C_P \|u_s^\varepsilon\|_V,$$

choosing γ such that $\gamma \max c_\xi C_P < 1$ we have:

$$\| \int_{t-\eta}^{t} c_\xi u_s^\varepsilon(\frac{i}{N}) e^{\Delta(t-s)} \delta_{\frac{i}{N}} \mathrm{d}s \|_V \leq \frac{1}{2|E|} \sup_{s\in[0,t]} \|u_s^\varepsilon\|_V.$$

Therefore, we have

$$\frac{1}{N} \sum_{\xi\in E} \sum_{i\in\mathcal{N}} \| \int_{t-\eta}^{t} c_\xi u_s^\varepsilon(\frac{i}{N}) e^{\Delta(t-s)} \delta_{\frac{i}{N}} \mathrm{d}s \|_V \leq \frac{1}{2} \sup_{s\in[0,t]} \|u_s^\varepsilon\|_V.$$

Now the second estimates of Lemma 2.2.1 gives:

$$\begin{aligned} \| \int_{0}^{t-\eta} c_\xi u_s^\varepsilon(\frac{i}{N}) e^{\Delta(t-s)} \delta_{\frac{i}{N}} \mathrm{d}s \|_V &\leq C_2(\eta) \max_{\xi\in E} |c_\xi| \int_0^{t-\eta} \|u_s^\varepsilon\|_\infty \mathrm{d}s \\ &\leq C_P C_2(\eta) \max_{\xi\in E} |c_\xi| \int_0^{t-\eta} \sup_{\sigma\in[0,s]} \|u_\sigma^\varepsilon\|_V \mathrm{d}s. \end{aligned}$$

Hence the existence of a constant α_2 depending only of T and η such that

$$\frac{1}{N} \sum_{\xi\in E} \sum_{i\in\mathcal{N}} \| \int_{0}^{t-\eta} c_\xi u_s^\varepsilon(\frac{i}{N}) e^{\Delta(t-s)} \delta_{\frac{i}{N}} \mathrm{d}s \|_V \leq \alpha_2 \int_0^t \sup_{\sigma\in[0,s]} \|u_\sigma^\varepsilon\|_V \mathrm{d}s.$$

Reassembling the above inequalities, we obtain

$$\sup_{s\in[0,t]} \|u_s^\varepsilon\|_V \leq \alpha_1 + \frac{1}{2} \sup_{s\in[0,t]} \|u_s^\varepsilon\|_V + \alpha_2 \int_0^t \sup_{\sigma\in[0,s]} \|u_\sigma^\varepsilon\|_V \mathrm{d}s.$$

The end of the proof is a classical application of the Gronwall's lemma. We find finally that $\sup_{s\in[0,t]} \|u_t^\varepsilon\|_V$ is bounded by a fixed constant independent of $t \in [0,T]$, ω and $\varepsilon \in]0,1]$. $\square$

Main assumptions and averaged equation.

We make the following assumptions. For any $y \in \mathbb{R}$ fixed, and any $j \in \{1, \cdots, l\}$, the fast generator $\mathcal{B}_j(y)$ is weakly irreducible on E_j, i.e. has a unique quasi-stationary distribution denoted by $\mu_j(y)$ (that is to say, $\mu_j(y)$ is nonnegative, $\mathcal{B}_j^*(y)\mu_j(y) = 0$ and we impose further that $\sum_{j=1}^{l} \mu_j(y)(E_j) = 1$). This quasi-stationary distribution is supposed to be, as is its derivative, Lipschitz-continuous in its argument y (see section 3.2.3 for more details).

Following [YZ98], the states in E_j can be considered as equivalent. For any $i \in \mathcal{N}$ we define the new stochastic process $(\bar{r}_t^\epsilon)_{t\geq 0}$ by $\bar{r}_t^\varepsilon(i) = j$ when $r_t^\varepsilon(i) \in E_j$ and abbreviate E_j by j . We then have an aggregate process $\bar{r}^\varepsilon(i)$ with values in $\{1, \cdots, l\}$. This is not a Markov process but we have the following proposition:

Proposition 3.2.3 ([YZ98]). *For any $y \in \mathbb{R}$, $i \in \mathcal{N}$ and $g : \{1, \cdots, l\} \to \mathbb{R}$, $\bar{r}^\varepsilon(i)$ converges weakly when ε goes to 0 to a Markov process $\bar{r}(i)$ generated by:*

$$\bar{\mathcal{B}}(y)g(\bar{r}(i)) = \sum_{j=1}^{l} 1_j(\bar{r}(i)) \sum_{k=1,k\neq j}^{l} (g(k) - g(j)) \sum_{\xi\in E_k} \sum_{\zeta\in E_j} \alpha_{\zeta,\xi}(y)\mu_j(\zeta). \tag{3.7}$$

The generator (3.7) expresses that if a state is in the class j, it jumps toward the class k at rate $\sum_{\xi\in E_k} \sum_{\zeta\in E_j} \alpha_{\zeta,\xi}(y)\mu_j(\zeta)$. That is, the rates of jumps of one state of E_j toward one state of E_k are averaged against the quasi-invariant measure associated to the class E_j. We then average the function $G_r(u)$ against the quasi-

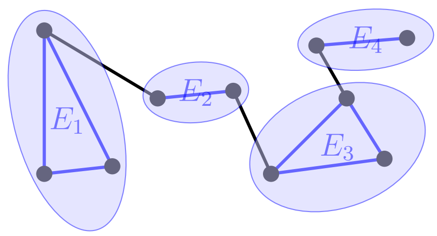


Figure 3.1: Aggregation of the states of E with $l = 4$.

invariant distributions. That is we consider that each pack of states E_j has reached its stationary behavior. For any $\bar{r} \in \bar{\mathcal{R}} = \{1, \cdots, l\}^{\mathcal{N}}$ we define the averaged function by:

$$F_{\bar{r}}(u) = \frac{1}{N} \sum_{i\in\mathcal{N}} \sum_{j=1}^{l} 1_j(\bar{r}(i)) \sum_{\zeta\in E_j} c_\zeta \mu_j(u(\frac{i}{N}))(\zeta)(v_\zeta - u(\frac{i}{N}))\delta_{\frac{i}{N}}. \tag{3.8}$$

Therefore, we call averaged equation of (3.2) the following PDE:

$$\partial_t u_t = \Delta u_t + F_{\bar{r}_t}(u_t) \tag{3.9}$$

with zero Dirichlet boundary condition and initial condition u_0 and $\bar{q}_0$ (where $\bar{q}_0$ is the aggregation of the initial channel configuration q_0). In equation (3.9), the process $(\bar{r}_t)_{t\in[0,T]}$ evolves, each coordinate independently over infinitesimal timescales, according to the averaged jump-rates between the subsets E_j of E.

Proposition 3.2.4. *For any $T > 0$, the equation (3.9) defines a PDMP $(u_t, \bar{r}_t)_{t\in[0,T]}$ in infinite dimensions in the sense of [BR11]. Moreover, there is a constant C such that:*

$$\sup_{t\in[0,T]} \|u_t\|_V \leq C$$

and $u \in \mathcal{C}([0,T], V)$ almost-surely.

Proof. The equation (3.9) is of the form (3.2) except that the indicator function is replaced by the probabilities μ_j. This changes nothing to the mathematics and the arguments developed in [Aus08] or [BR11] still apply. We refer the reader to [Aus08] and [BR11] for more details. □

We can now state our main result which states that our approximation (3.9) is valid.

Theorem 3.2.2. *The stochastic process u^ε solution of (3.2) converges in law to the solution of (3.9) when ε goes to 0.*

More precisely, we will prove the following proposition:

Proposition 3.2.5. *$(u^\varepsilon, \varepsilon \in]0,1])$ is tight in $\mathcal{C}([0,T], V)$ and any accumulation point u verifies:*

$$(u_t, \phi)_{L^2(I)} = (u_0, \phi)_{L^2(I)} + \int_0^t (u_s, \phi'')_{L^2(I)} ds + \int_0^t < F_{\bar{r}_s}(u_s), \phi > ds \qquad (3.10)$$

for all ϕ in $\mathcal{C}_0^2(I)$. Moreover the accumulation point is unique up to indistinguishability.

Plan of the campaign.

To make the proof of our main result easier to read, we will consider in a first step that all the states communicate at a fast rate. This is called the all-fast case. We will prove our main result in detail in this case and then give the key points for the validity of our proof in the general case. In the all-fast case there is a unique class E and the generator has the simplest form:

$$\mathcal{B}f(u,r) = \sum_{i\in\mathcal{N}} \sum_{\zeta\in E} [f(u, r(r(i) \to \zeta)) - f(u,r)] \alpha_{r(i),\zeta}(u(\frac{i}{N})). \qquad (3.11)$$

We make the following assumptions: for any $i \in \mathcal{N}$ with $u \in V$ held fixed, the Markov process $r(i)$ has a unique stationary distribution $\mu(u(\frac{i}{N}))$. Then the process $(r(i), i \in \mathcal{N})$ has the following stationary distribution:

$$\mu(u) = \bigotimes_{i \in \mathcal{N}} \mu(u(\frac{i}{N})).$$

We assume that the rate functions are Lipschitz continuous. The average (generalized) function is then:

$$\begin{aligned} F(u) &= \int_{\mathcal{R}} G_r(u)\mu(u)(dr) \qquad (3.12) \\ &= \frac{1}{N}\sum_{\xi \in E}\sum_{i \in \mathcal{N}} c_\xi \mu(u(\frac{i}{N}))(\xi)(v_\xi - u(\frac{i}{N}))\delta_{\frac{i}{N}}, \end{aligned}$$

if u is held fixed.

Is this hypothesis reasonable? For the most classical model in neurosciences, the rate functions are indeed, as well as their derivatives, Lipschitz continuous. See for example the rate functions in classical Hodgkin-Huxley model [HH52] (see also the Appendix 3.A).

Let us show briefly that in this case the ergodic measure μ is Lipschitz continuous in its argument. Let i be in $\mathcal{N}$ and the membrane potential at $\frac{i}{N}$ be fixed equal to $x \in \mathbb{R}$, the state space of the continuous time Markov chain $r(i)$ is E which is a finite set. We set $|E| = m$. We can then enumerate the elements of E : $E = \{\xi_1, \cdots, \xi_m\}$. The generator of $r(i)$ is then the matrix :

$$J(x) = (\alpha_{\xi_i,\xi_j}(x))_{1 \le i,j \le m}$$

with $\alpha_{\xi_i,\xi_i}(x) = -\sum_{j \neq i} \alpha_{\xi_i,\xi_j}(x)$. Therefore, if we assumed that our continuous time Markov chain is ergodic, it has an invariant measure which is of the form:

$$\mu(x)(\xi) = \frac{\sum \prod \alpha_{\xi,\xi'}(x)}{\sum \prod \alpha_{\zeta,\zeta'}(x)}.$$

The sums and products involved here are calculated over subsets of E which are given by resolving the equation $\mu(x)^T J = 0$. μ is then Lipschitz continuous in x since each $\alpha_{\xi,\zeta}$ is Lipschitz continuous in x and bounded.

The assumption of independence of each coordinate of the process r over infinitesimal timescales is important to have a simple expression of the invariant measure of the process r and simplify the mathematics. It is possibly unrealistic from the biological point of view but is often assumed in mathematical neurosciences.

The plan of the campaign to prove our main result (3.2.2) in the all-fast case is the following. We want to prove the convergence in law of a family of càdlàg stochastic processes with values in a Hilbert space towards another, here deterministic, process. There is a well known strategy to do that:

- prove the tightness of the family
- identify the limit

In our case, tightness (see subsection 3.3.1) will follow roughly from the uniform boundedness of the family $(u^\varepsilon, \varepsilon \in]0,1])$ obtained in Proposition 3.2.2. The identification of the limit will follow on one side from tightness and on the other side from some classical argument in the theory of averaging for stochastic processes, see for example [PS08]. We will introduce a Poisson equation which will enable us to control some problematic terms (see Proposition 3.3.1) and identify the limit (subsection 3.3.1). We will then move to the general case with multiple classes (in subsection 3.3.2) with an emphasis on each point which could give an issue and show how things work in this case.

3.3 Proof of the main result

3.3.1 The all-fast case

We want to prove Proposition 3.2.5. Let us recall that in the all-fast case $F_{\bar{r}}$ in (3.10) reduces to F in (3.12).

Let us outline the strategy of the proof following the plan of the campaign announced in Section 3.2.3.

- Tightness.
  1. Thanks to Proposition 3.2.2, use the Markov inequality to prove that Aldous's condition and property (2.2) of Theorems 2.3.3 and 2.3.2 hold.
  2. For $t \in [0,T]$ fixed, find the finite dimensional space $H_{\delta,\eta}$ of (2.3), Theorem 2.3.2 using a truncation of the solution u_t^ε in the mild formulation. Bound $\mathbb{E}(d(u_t^\varepsilon, H_{\delta,\eta}) > \eta)$ and use the Markov inequality.
  3. Obtain tightness in $\mathbb{D}([0,T],V)$ by Theorems 2.3.3 and 2.3.2 and conclude that tightness in $\mathcal{C}([0,T],V)$ holds as well thanks to the regularity of the solution u^ε for any $\varepsilon \in]0,1]$.
- Identification of the limit.

1. The aim of this part of the proof is to show that for any $\phi \in \mathcal{C}_0^2(I)$:
$$(u_t^\varepsilon, \phi)_{L^2(I)} = (u_0, \phi)_{L^2(I)} + \int_0^t (u_s^\varepsilon, \phi'')_{L^2(I)} \mathrm{d}s + \int_0^t < F(u_s^\varepsilon), \phi > \mathrm{d}s + \gamma(\varepsilon),$$
where $\gamma(\varepsilon)$ goes to zero with ε.
2. Begin by showing that the following property is sufficient:
$$\lim_{\varepsilon \to 0} \mathbb{E} \left| \int_0^t < G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), \phi > \mathrm{d}s \right| = 0 \quad \forall \phi \in \mathcal{C}_0^2(I).$$
To prove this property, use the tightness of the family $(u^\varepsilon, \varepsilon \in]0,1])$.
3. Obtain the decomposition $\int_0^t < G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), \phi > \mathrm{d}s = a^{t,\varepsilon,\eta} + \int_0^t (G^\eta_{r_s^\varepsilon}(u_s^\varepsilon) - F^\eta(u_s^\varepsilon), \phi)_{L^2(I)} \mathrm{d}s + c^{t,\varepsilon,\eta}$ where for $\eta > 0$, G^η and F^η are smoother versions of G and F respectively. More precisely:
$$\begin{aligned} a_{t,\varepsilon,\eta} &= \int_0^t < G_{r_s^\varepsilon}(u_s^\varepsilon) - G^\eta_{r_s^\varepsilon}(u_s^\varepsilon), \phi > \mathrm{d}s, \\ c_{t,\varepsilon,\eta} &= \int_0^t < F(u_s^\varepsilon) - F^\eta(u_s^\varepsilon), \phi > \mathrm{d}s. \end{aligned}$$
4. Show that $\lim_{\eta \to 0} \lim_{\varepsilon \to 0} |a^{t,\varepsilon,\eta}| + |c^{t,\varepsilon,\eta}| = 0$.
5. Replace the term $(G^\eta_{r_s^\varepsilon}(u_s^\varepsilon) - F^\eta(u_s^\varepsilon), \phi)_{L^2(I)}$ by $\mathcal{B}^\eta f^\eta(u_s^\varepsilon, r_s^\varepsilon)$ where f^η is the solution of a Poisson equation and $\mathcal{B}^\eta$ is a smoother version of $\mathcal{B}$. Show that this solution exists and has some nice properties. Show that $\lim_{\eta \to 0} \lim_{\varepsilon \to 0} \mathbb{E} \left| \int_0^t \mathcal{B}^\eta f^\eta(u_s^\varepsilon, r_s^\varepsilon) - \mathcal{B} f^\eta(u_s^\varepsilon, r_s^\varepsilon) \mathrm{d}s \right| = 0$.
6. Write the semi-martingale expansion of $\int_0^t \mathcal{B} f^\eta(u_s^\varepsilon, r_s^\varepsilon) \mathrm{d}s$. Bound its martingale part thanks to the bound for the bracket of the martingale.
7. Bound its finite variation part thanks to the nice properties of f^η and the uniform bound from Proposition 3.2.2.
8. Aggregate all these bounds together to obtain that
$$\mathbb{E} \left| \int_0^t \mathcal{B} f^\eta(u_s^\varepsilon, r_s^\varepsilon) \mathrm{d}s \right|^2$$
is bounded by a constant (which depends on η) times ε.
9. Let ε goes to zero and then η goes to zero and conclude that (3.10) holds and implies uniqueness up to indistinguishability.

Tightness. STEP 1 (Property (2.2) and Aldous condition of Theorem 2.2). We begin with the Aldous condition. Let τ be a stopping time (for the natural filtration associated to the process $(u^\varepsilon, r^\varepsilon)$) and θ such that $\tau+\theta<T$. We need to control the size of $\|u^\varepsilon_{\tau+\theta}-u^\varepsilon_\tau\|_V$. At first we work deterministically (or "ω by ω"). We break the difference $\|u^\varepsilon_{\tau+\theta}-u^\varepsilon_\tau\|^2_V$ into three pieces using the mild formulation of u^ε in (3.6):

$$\begin{aligned}
&\|u^\varepsilon_{\tau+\theta}-u^\varepsilon_\tau\|^2_V\\
\leq\ & 4\|e^{\Delta(\tau+\theta)}u_0-e^{\Delta\tau}u_0\|^2_V+4\|\int_\tau^{\tau+\theta}e^{\Delta(\tau+\theta-s)}G_{r^\varepsilon_s}(u^\varepsilon_s)\mathrm{d}s\|^2_V\\
+\ & 4\|\int_0^\tau\left(e^{\Delta(\tau+\theta-s)}-e^{\Delta(\tau-s)}\right)G_{r^\varepsilon_s}(u^\varepsilon_s)\mathrm{d}s\|^2_V.
\end{aligned}$$

Let us show now that the supremum over $\theta\in]0,\epsilon[$ of each term on the right hand side of the above inequality goes to 0 with ϵ (we take $\epsilon>0$). We start with the third term:

$$\|\int_0^\tau\left(e^{\Delta(\tau+\theta-s)}-e^{\Delta(\tau-s)}\right)G_{r^\varepsilon_s}(u^\varepsilon_s)\mathrm{d}s\|^2_V.$$

Notice that:

$$\begin{aligned}
&\left(e^{\Delta(\tau+\theta-s)}-e^{\Delta(\tau-s)}\right)G_{r^\varepsilon_s}(u^\varepsilon_s)\\
=\ &\sum_{k\geq1}e^{-(k\pi)^2(\tau-s)}\left(e^{-(k\pi)^2\theta}-1\right)\frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}}g_\xi(s,i)(1+(k\pi)^2)e_k\left(\frac{i}{N}\right)e_k,
\end{aligned}$$

where we have used the fact that $<e_k,\delta_{\frac{i}{N}}>=(1+(k\pi)^2)e_k(\frac{i}{N})$ (recall that $e_k=\frac{\sqrt{2}}{\sqrt{1+(k\pi)^2}}\sin(k\pi\cdot)$) and denoted $c_\xi 1_\xi(r^\varepsilon_s(i))(v_\xi-u^\varepsilon_s(\frac{i}{N}))$ by $g_\xi(s,i)$. Thus we obtain:

$$\begin{aligned}
&\|\int_0^\tau\left(e^{\Delta(\tau+\theta-s)}-e^{\Delta(\tau-s)}\right)G_{r^\varepsilon_s}(u^\varepsilon_s)\mathrm{d}s\|^2_V\\
&\leq 4|E|^2\max_{\xi\in E}c_\xi^2(\max_{\xi\in E}v_\xi^2+C^2)\sum_{k\geq1}\left(\int_0^\tau e^{-(k\pi)^2(\tau-s)}\left(e^{-(k\pi)^2\theta}-1\right)\sqrt{1+(k\pi)^2}\mathrm{d}s\right)^2\\
&=4|E|^2\max_{\xi\in E}c_\xi^2(\max_{\xi\in E}v_\xi^2+C^2)\sum_{k\geq1}\left(e^{-(k\pi)^2\theta}-1\right)^2(1+(k\pi)^2)\frac{\left(1-e^{-(k\pi)^2\tau}\right)^2}{(k\pi)^4}\\
&\leq 4|E|^2\max_{\xi\in E}c_\xi^2(\max_{\xi\in E}v_\xi^2+C^2)\sum_{k\geq1}\left(e^{-(k\pi)^2\theta}-1\right)^2\frac{1+(k\pi)^2}{(k\pi)^4},
\end{aligned}$$

where C is the constant of Proposition 3.2.2. By dominated convergence, we see that:

$$\lim_{\eta\to 0}\sup_{\theta\in]0,\epsilon[}\sum_{k\geq 1}\left(e^{-(k\pi)^2\theta}-1\right)^2\frac{1+(k\pi)^2}{(k\pi)^4}\leq\lim_{\eta\to 0}\sum_{k\geq 1}\left(e^{-(k\pi)^2\epsilon}-1\right)^2\frac{1+(k\pi)^2}{(k\pi)^4}=0$$

For the second term, as before, we can show that:

$$\begin{aligned}
& \|\int_\tau^{\tau+\theta} e^{\Delta(\tau+\theta-s)}G_{r^\varepsilon_s}(u^\varepsilon_s)\mathrm{d}s\|_V^2 \\
\leq\ & 4|E|^2\max_{\xi\in E}c_\xi^2(\max_{\xi\in E}v_\xi^2+C^2)\sum_{k\geq 1}\left(\int_\tau^{\tau+\theta} e^{-(k\pi)^2(\tau+\theta-s)}\sqrt{1+(k\pi)^2}\mathrm{d}s\right)^2 \\
\leq\ & 4|E|^2\max_{\xi\in E}c_\xi^2(\max_{\xi\in E}v_\xi^2+C^2)\sum_{k\geq 1}(1+(k\pi)^2)\frac{\left(1-e^{-(k\pi)^2\theta}\right)^2}{(k\pi)^4}.
\end{aligned}$$

By dominated convergence, we see that:

$$\lim_{\epsilon\to 0}\sup_{\theta\in]0,\epsilon[}\sum_{k\geq 1}(1+(k\pi)^2)\frac{\left(1-e^{-(k\pi)^2\theta}\right)^2}{(k\pi)^4}\leq\lim_{\eta\to 0}\sum_{k\geq 1}(1+(k\pi)^2)\frac{\left(1-e^{-(k\pi)^2\epsilon}\right)^2}{(k\pi)^4}=0.$$

For the first term we have, by the Bessel-Parseval equality:

$$\begin{aligned}
& \|e^{\Delta(\tau+\theta)}u_0-e^{\Delta\tau}u_0\|_V^2 \\
=\ & \|\sum_{k\geq 1}e^{-(k\pi)^2\tau}(e^{-(k\pi)^2\theta}-1)(u_0,e_k)e_k\|_V^2 \\
=\ & \sum_{k\geq 1}e^{-2(k\pi)^2\tau}(e^{-(k\pi)^2\theta}-1)^2(u_0,e_k)^2 \\
\leq\ & \sum_{k\geq 1}(e^{-(k\pi)^2\theta}-1)^2(u_0,e_k)^2.
\end{aligned}$$

By dominated convergence, we see that:

$$\lim_{\epsilon\to 0}\sup_{\theta\in]0,\epsilon[}\sum_{k\geq 1}(e^{-(k\pi)^2\theta}-1)^2(u_0,e_k)^2\leq\lim_{\eta\to 0}\sum_{k\geq 1}(e^{-(k\pi)^2\epsilon}-1)^2(u_0,e_k)^2=0$$

Combining the results obtained for the three terms we have:

$$\lim_{\eta\to 0}\sup_{\theta\in]0,\epsilon[}\|u^\varepsilon_{\tau+\theta}-u^\varepsilon_\tau\|_V^2=0$$

uniformly in $\varepsilon \in]0,1]$. Therefore, for all $M, \delta > 0$, using the Chebyshev inequality, we can choose η so small that:

$$\sup_{\varepsilon\in]0,1]}\sup_{\theta\in]0,\epsilon[}\mathbb{P}(\|u^\varepsilon_{\tau+\theta}-u^\varepsilon_\tau\|_V \geq M) \leq \sup_{\varepsilon\in]0,1]}\sup_{\theta\in]0,\epsilon[}\frac{\mathbb{E}(\|u^\varepsilon_{\tau+\theta}-u^\varepsilon_\tau\|^2_V)}{M^2} < \delta$$

The first condition (2.2) of Theorem 2.3.2 is verified by the application of the Markov inequality. Indeed, for any $t \in [0,T]$ and $\delta > 0$ by Proposition 3.2.2 there exists a constant $C > 0$ independent of $\varepsilon \in]0,1]$ and $t \in [0,T]$ such that:

$$\sup_{\varepsilon\in]0,1]}\mathbb{P}(\|u^\varepsilon_t\|_V > \rho) \leq \frac{1}{\rho}\sup_{\varepsilon\in]0,1]}\mathbb{E}\|u^\varepsilon_t\|_V \leq \frac{C}{\rho} < \delta$$

for any $\rho > 0$ large enough.

STEP 2 (Truncation). We have to show that for any $\delta, \epsilon > 0$ we can find $\varepsilon_0 > 0$ and a space $H_{\delta,\epsilon}$ such that:

$$\sup_{\varepsilon\in]0,\varepsilon_0]}\mathbb{P}(d(u^\varepsilon_t, H_{\delta,\epsilon}) > \epsilon) \leq \delta,$$

where $d(u^\varepsilon_t, H_{\delta,\epsilon}) = \inf_{v\in H_{\delta,\epsilon}}\|u^\varepsilon_t - v\|_V$. We fix $t \in [0,T]$. Recalling the mild representation (3.6), we have:

$$u^\varepsilon_t = e^{\Delta t}u_0 + \int_0^t \frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}} c_\xi 1_\xi(r^\varepsilon_s(i))(v_\xi - u^\varepsilon_s(\frac{i}{N}))e^{\Delta(t-s)}\delta_{\frac{i}{N}}\mathrm{d}s,$$

where, using the explicit expression of the semi-group $e^{\Delta t}$, $e^{\Delta(t-s)}\delta_{\frac{i}{N}}$ is equal to

$$\sum_{k\geq 1} e^{-(k\pi)^2(t-s)}(1+(k\pi)^2)e_k(\frac{i}{N})e_k.$$

We define, for $f \in V$ and $x \in I$ the truncations up to the order p of u^ε_t and $e^{\Delta t}$:

$$\begin{aligned} e^{\Delta t}_p f &:= \sum_{k=1}^p e^{-(k\pi)^2 t}(f,e_k)e_k \text{ and } e^{\Delta t}_p\delta_x := \sum_{k=1}^p e^{-(k\pi)^2 t}(1+(k\pi)^2)e_k(x)e_k, \\ u^{\varepsilon,p}_t &:= e^{\Delta t}_p u_0 + \int_0^t \frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}} c_\xi 1_\xi(r^\varepsilon_s(i))(v_\xi - u^\varepsilon_s(\frac{i}{N}))e^{\Delta(t-s)}_p\delta_{\frac{i}{N}}\mathrm{d}s. \end{aligned}$$

For any $\delta, \epsilon > 0$ we can find p independent of ε such that $\|u^\varepsilon_t - u^{\varepsilon,p}_t\|_V \leq C'\epsilon\delta$ with the constant C' deterministic and independent of δ, ϵ and $\varepsilon \in]0,1]$. Indeed we can easily show, as in STEP 1, that:

$$\|u^\varepsilon_t - u^{\varepsilon,p}_t\|^2_V \leq 2\|\sum_{k=p+1}^\infty e^{-(k\pi)^2 t}(u_0,e_k)e_k\|^2_V + 2C'\sum_{k=p+1}^\infty (1+(k\pi)^2)\frac{(1-e^{-(k\pi)^2 t})^2}{(k\pi)^4},$$

which is independent of $\varepsilon \in]0,1]$. The convergence of each series, uniformly in ε, enables us to choose a suitable p independent of ε. Let us denote $V_p = \text{span}\{e_i, 1 \leq i \leq p\}$. We choose $H_{\delta,\epsilon} = V_p$. Since $u_t^{\varepsilon,p} \in H_{\delta,\epsilon} = V_p$

$$\mathbb{E}(d(u_t^\varepsilon, H_{\delta,\epsilon})) \leq \mathbb{E}(\|u_t^\varepsilon - u_t^{\varepsilon,p}\|_V) \leq C'\epsilon\delta.$$

The Markov's inequality gives us :

$$\mathbb{P}(d(u_t^\varepsilon, H_{\delta,\epsilon}) > \epsilon) \leq C'\delta$$

with δ independent of $\varepsilon \in]0,1]$.

STEP 3 (Tightness). Therefore using Theorems 2.3.2 and 2.3.3, the family $\{u^\varepsilon, \varepsilon \in]0,1]\}$ is tight in $\mathbb{D}([0,T],V)$. Since we know that for each $\varepsilon \in]0,1]$, u^ε is in $\mathcal{C}([0,T],V)$ we have in fact that $\{u^\varepsilon, \varepsilon \in]0,1]\}$ is tight in $\mathcal{C}([0,T],V)$ (the Skorokhod topology restrict to continuous functions coincides with the uniform topology).

□

Identification of the limit. STEP 1 (Reduction of the problem). We know that the family $(u^\varepsilon, \varepsilon \in]0,1])$ is tight in $\mathcal{C}([0,T],V)$. We still denote by u^ε a converging sub-sequence of u^ε and we denote by u the corresponding accumulation point. We want to show that for all ϕ in $\mathcal{C}_0^2(I)$:

$$(u_t,\phi)_{L^2(I)} = (u_0,\phi)_{L^2(I)} + \int_0^t (u_s,\phi'')_{L^2(I)}\mathrm{d}s + \int_0^t < F(u_s),\phi > \mathrm{d}s,$$

almost surely. We will show that:

$$\mathbb{E}\left|(u_t,\phi)_{L^2(I)} - (u_0,\phi)_{L^2(I)} - \int_0^t (u_s,\phi'')_{L^2(I)}\mathrm{d}s - \int_0^t < F(u_s),\phi > \mathrm{d}s\right| = 0. \tag{3.13}$$

Notice that:

$$\begin{aligned} & \mathbb{E}\left|(u_t^\varepsilon,\phi)_{L^2(I)} - (u_0,\phi)_{L^2(I)} - \int_0^t (u_s^\varepsilon,\phi'')_{L^2(I)}\mathrm{d}s - \int_0^t < F(u_s^\varepsilon),\phi > \mathrm{d}s\right| \\ = \; & \mathbb{E}\left|\int_0^t < G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon),\phi > \mathrm{d}s\right|. \end{aligned}$$

Since the function:

$$h(v) = \left|(v_t,\phi)_{L^2(I)} - (v_0,\phi)_{L^2(I)} - \int_0^t (v_s,\phi'')_{L^2(I)}ds - \int_0^t < F(v_s),\phi > \mathrm{d}s\right|$$

is continuous on $\mathcal{C}([0,T],V)$ and the arguments $(u^\varepsilon, \varepsilon \in]0,1])$ and u are uniformly bounded in ε in $\mathcal{C}([0,T],V)$, by convergence in law we have that:

$$\begin{aligned}\lim_{\varepsilon\to 0} \quad & \mathbb{E}\left|(u_t^\varepsilon,\phi)_{L^2(I)} - (u_0,\phi)_{L^2(I)} - \int_0^t (u_s^\varepsilon,\phi'')_{L^2(I)}\mathrm{d}s - \int_0^t < F(u_s^\varepsilon),\phi > \mathrm{d}s\right| \\ = \quad & \mathbb{E}\left|(u_t,\phi)_{L^2(I)} - (u_0,\phi)_{L^2(I)} - \int_0^t (u_s,\phi'')_{L^2(I)}\mathrm{d}s - \int_0^t < F(u_s),\phi > \mathrm{d}s\right|.\end{aligned}$$

Therefore (3.13) will follow if $\lim_{\varepsilon\to 0}\mathbb{E}\left|\int_0^t < G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon),\phi > \mathrm{d}s\right| = 0$.

STEP 2 (Mollification). Let $\eta > 0$. For $i \in \mathcal{N}$ define the mollifier $\psi_i^\eta \in \mathcal{C}^\infty(I,\mathbb{R})$ such that

$$\forall \phi \in \mathcal{C}(I,\mathbb{R}), \quad \lim_{\eta\to 0}(\psi_i^\eta,\phi)_{L^2(I)} = \phi(\frac{i}{N}).$$

We define mollifications G^η, F^η, $\mathcal{B}^\eta$ and μ^η of the reaction terms G and F, the operator $\mathcal{B}$ and the measure μ respectively. Let $f : L^2(I)\times\mathcal{R}\to\mathbb{R}$ be measurable and bounded on $\mathcal{R}$ and continuous and bounded on $L^2(I)$. For $u \in L^2(I)$ and $r \in \mathcal{R}$ we define

$$G_r^\eta(u) = \frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}} c_\xi 1_\xi(r(i))(v_\xi - (u,\psi_i^\eta)_{L^2(I)})\psi_i^\eta, \tag{3.14}$$

$$F^\eta(u) = \sum_{\xi\in E} c_\xi \mu((u,\psi_i^\eta)_{L^2(I)})(\xi)(v_\xi - (u,\psi_i^\eta)_{L^2(I)})\psi_i^\eta, \tag{3.15}$$

$$\mathcal{B}^\eta f(u,r) = \sum_{i\in\mathcal{N}}\sum_{\zeta\in E}[f(u,r(r(i)\to\zeta)) - f(u,r)]\alpha_{r(i),\zeta}((u,\psi_i^\eta)_{L^2(I)}), \tag{3.16}$$

$$\mu^\eta(u) = \bigotimes_{i\in\mathcal{N}}\mu((u,\psi_i^\eta)_{L^2(I)}). \tag{3.17}$$

We use the following decomposition of $\int_0^t < G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon),\phi > \mathrm{d}s$:

$$\begin{aligned}\int_0^t < G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon),\phi > \mathrm{d}s = & \int_0^t < G_{r_s^\varepsilon}(u_s^\varepsilon) - G^\eta_{r_s^\varepsilon}(u_s^\varepsilon),\phi > \mathrm{d}s \quad (a_{t,\varepsilon,\eta}) \\ & + \int_0^t (G^\eta_{r_s^\varepsilon}(u_s^\varepsilon) - F^\eta(u_s^\varepsilon),\phi)_{L^2(I)}\mathrm{d}s \quad (b_{t,\varepsilon,\eta}) \\ & + \int_0^t < F^\eta(u_s^\varepsilon) - F(u_s^\varepsilon),\phi > \mathrm{d}s \quad (c_{t,\varepsilon,\eta}).\end{aligned}$$

STEP 3 (Bound $a_{t,\varepsilon,\eta}$ and $c_{t,\varepsilon,\eta}$). Note that

$$a_{t,\varepsilon,\eta} = \frac{1}{N}\sum_{i\in\mathcal{N}}\int_0^t c_{r_s^\varepsilon(i)}[(v_{r_s^\varepsilon(i)} - u_s^\varepsilon(\frac{i}{N}))\phi(\frac{i}{N}) - (v_{r_s^\varepsilon(i)} - (u_s^\varepsilon,\psi_i^\eta)_{L^2(I)})(\psi_i^\eta,\phi)_{L^2(I)}]\mathrm{d}s.$$

A straightforward calculation leads to the estimate

$$\begin{aligned}|a_{t,\varepsilon,\eta}| \leq T \max_{\xi\in E}(|c_\xi v_\xi| + |c_\xi| \max_{s\in[0,T]} \max_{x\in I} |u^\varepsilon_s(x)|) \max_{i\in\mathcal{N}} |(\psi^\eta_i, \phi)_{L^2(I)} - \phi(\frac{i}{N})| \\ + \max_{\xi\in E} |c_\xi| \max_{i\in\mathcal{N}} |(\psi^\eta_i, \phi)_{L^2(I)}| \int_0^T \max_{i\in\mathcal{N}} |u^\varepsilon_s(\frac{i}{N}) - (\psi^\eta_i, u^\varepsilon_s)_{L^2(I)}| \mathrm{d}s.\end{aligned}$$

Using the property of the mollifiers ψ^η_i, the uniform boundedness of the family $\{u^\varepsilon, \varepsilon \in]0,1]\}$ in V and the convergence in law of u^ε towards u in $\mathcal{C}([0,T], V)$, it is not difficult to see that:

$$\lim_{\eta\to 0} \lim_{\varepsilon\to 0} \mathbb{E}(|a_{t,\varepsilon,\eta}|) = 0.$$

Similarly one can show that

$$\lim_{\eta\to 0} \lim_{\varepsilon\to 0} \mathbb{E}(|c_{t,\varepsilon,\eta}|) = 0.$$

The remainder of the proof is therefore devoted to show that:

$$\lim_{\eta\to 0} \lim_{\varepsilon\to 0} \mathbb{E} \left| \int_0^t (G^\eta_{r^\varepsilon_s}(u^\varepsilon_s) - F^\eta(u^\varepsilon_s), \phi)_{L^2(I)} \mathrm{d}s \right| = 0.$$

STEP 4 (Poisson equation). We introduce the following Poisson equation on $f : L^2(I) \times \mathcal{R} \to \mathbb{R}$:

$$\begin{aligned}\mathcal{B}^\eta f(u,r) &= (G^\eta_r(u) - F^\eta(u), \phi)_{L^2(I)} \\ \int_{\mathcal{R}} f(u,r)\mu^\eta(u)(dr) &= 0\end{aligned} \tag{3.18}$$

with $\mathcal{B}^\eta$ the generator given by equation (3.16):

$$\mathcal{B}^\eta f(u,r) = \sum_{i\in\mathcal{N}} \sum_{\zeta\in E} [f(u, r(r(i) \to \zeta)) - f(u,r)] \alpha_{r(i),\zeta}((u, \phi^\eta_i)_{L^2(I)})$$

and μ^η given by equation (3.17). Recall also that we identify the dual of $L^2(I)$ with itself.

Proposition 3.3.1. *The Poisson equation (3.18) has a unique solution f^η which is measurable and locally Lipschitz continuous in its first variable with respect to the $\|\cdot\|_{L^2(I)}$ norm. Moreover $\sup_{s\in[0,T]} |f^\eta(u^\varepsilon_s, r^\varepsilon_s)|$ is bounded almost surely by a constants independent of ε and η.*
For all fixed $r \in \mathcal{R}$, the map $u \in H \mapsto f^\eta(u,r)$ has a Fréchet derivative denoted by $\frac{\mathrm{d}f^\eta}{\mathrm{d}u}(u,r)$. For $(u,r) \in L^2(I) \times \mathcal{R}$, one may identify the linear operator $\frac{\mathrm{d}f^\eta}{\mathrm{d}u}(u,r)$ on $L^2(I)$ with an element $f^\eta_u(u,r)$ of $L^2(I)$ which is also an element of V and is bounded in V almost surely by a constant independent of ε (but which depends on η).

We begin the proof of the above Proposition by the following Lemma.

Lemma 3.3.1. *Let $B \subset L^2(I)$ be a bounded domain and r be in $\mathcal{R}$. The map $u \in L^2(I) \mapsto (G^\eta_r(u) - F^\eta(u), \phi)_{L^2(I)}$ is:*

1. *bounded on B by a constant independent of r and η.*
2. *Lipschitz continuous on B with Lipschitz constant independent of r.*
3. *Fréchet differentiable on B with Fréchet derivative in $u \in H$ denoted by $\frac{\mathrm{d}f}{\mathrm{d}u}(u, r)$. Moreover the Fréchet derivative is bounded in $\|\cdot\|_{L^2(I)}$ norm uniformly in $u \in B$ and $r \in \mathcal{R}$ (but not uniformly in η).*

Proof. Recall that:

$$\begin{aligned}&(G^\eta_r(u) - F^\eta(u), \phi)_{L^2(I)}\\ &= \frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}} c_\xi(1_\xi(r(i)) - \mu((u, \psi^\eta_i)_{L^2(I)})(\xi))(v_\xi - (u, \psi^\eta_i)_{L^2(I)})(\psi^\eta_i, \phi)_{L^2(I)}.\end{aligned}$$

Since for $\xi \in E$, $\mu(\cdot)(\xi)$ is bounded and Lipschitz continuous on $\mathbb{R}$, the two first points follow. For the third point we note that the map $u \in H \mapsto (G_r(u) - F(u), \phi)_{L^2(I)}$ is a linear combination of the Fréchet differentiable functions $u \in H \mapsto \mu\left((u, \psi^\eta_i)_{L^2(I)}\right)(\xi)$, $u \mapsto (u, \psi^\eta_i)_{L^2(I)}$ and of the product of these two functions. The Fréchet derivative in $u \in L^2(I)$ against $h \in L^2(I)$ is given by:

$$\begin{aligned}&-\frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}} c_\xi[\mu'((u, \psi^\eta_i)_{L^2(I)})(\xi)(v_\xi - (u, \psi^\eta_i)_{L^2(I)}) + (1_\xi(r(i)) - \mu((u, \psi^\eta_i)_{L^2(I)})(\xi))]\\ &\qquad\qquad \times (\phi, \psi^\eta_i)_{L^2(I)}(h, \psi^\eta_i)_{L^2(I)},\end{aligned}$$

which gives us the boundedness property uniformly in $r \in \mathcal{R}$ and $u \in B$. Remark also that one can identify the Fréchet derivative, which is an element of $(L^2(I))^*$ with the element

$$\begin{aligned}&-\frac{1}{N}\sum_{\xi\in E}\sum_{i\in\mathcal{N}} c_\xi[\mu'((u, \psi^\eta_i)_{L^2(I)})(\xi)(v_\xi - (u, \psi^\eta_i)_{L^2(I)}) + (1_\xi(r(i)) - \mu((u, \psi^\eta_i)_{L^2(I)})(\xi))]\\ &\qquad\qquad \times (\phi, \psi^\eta_i)_{L^2(I)}\psi^\eta_i\end{aligned}$$

of $L^2(I)$ which is also an element of V. Note that this element is bounded in V for bounded arguments, but not uniformly in η. □

Proof of Proposition 3.3.1. For $u \in L^2(I)$ held fixed, $\mathcal{B}^\eta(u, \cdot)$ is an operator on $\mathbb{R}^{\mathcal{R}}$ which is a space of finite dimension. The Fredholm alternative in such a space is

$$\mathrm{Im}(\mathcal{B}^\eta) = (\ker((\mathcal{B}^\eta)^*))^\perp.$$

Moreover, by the ergodicity assumption, for any fixed $u \in L^2(I)$, $\ker((\mathcal{B}^\eta)^*(u,\cdot)) = \mathrm{span}(\mu^\eta(u))$. Therefore equation (3.18) with $u \in L^2(I)$ fixed has a solution if and only if:

$$\int_{\mathcal{R}} \mu^\eta(u)(\mathrm{d}r)(G_r^\eta(u) - F^\eta(u), \phi)_{L^2(I)} = 0.$$

This latter equality holds by the definition of F^η and the bi-linearity of the scalar product $(\cdot,\cdot)_{L^2(I)}$. Moreover, we can then always choose f^η satisfying $\int_{\mathcal{R}} f^\eta(u,r)\mu^\eta(u)(dr) = 0$ by taking its projection on $(\ker((\mathcal{B}^\eta)^*))^\perp$. Thus we have a solution $f^\eta(u,\cdot)$ for any $u \in L^2(I)$ fixed. Uniqueness of f^η follows easily from the condition

$$\int_{\mathcal{R}} f^\eta(u,r)\mu^\eta(u)(\mathrm{d}r) = 0.$$

Recall that for all $\varepsilon \in]0,1]$, $u^\varepsilon \in B$ where $B = \{u \in L^2(I); \|u\|_{L^2(I)} \leq C\}$

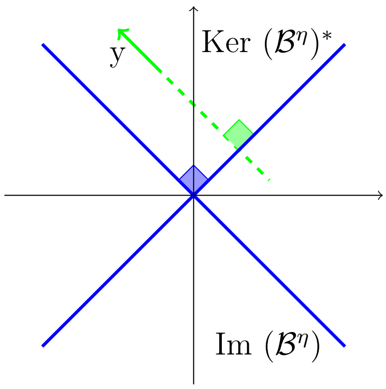


Figure 3.2: Fredholm alternative in finite dimension.

where C is the deterministic constant given by Proposition 3.2.2. Then the desired properties of f^η follow from Lemma 3.3.1 and the fact that for the bounded domain $B \subset L^2(I)$, each function $u \in B \mapsto \alpha_{r(i),\zeta}((u,\psi_i^\eta)_{L^2(I)})$ is bounded below and above by strictly positive constants, Lipschitz continuous and Fréchet differentiable with Fréchet derivative uniformly bounded in u, ζ and r. □

We thus have:

$$\begin{aligned}
&\int_0^t (G^\eta_{r_s^\varepsilon}(u_s^\varepsilon) - F^\eta(u_s^\varepsilon), \phi)_{L^2(I)}\mathrm{d}s \\
&= \int_0^t \mathcal{B}^\eta f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s \\
&= \int_0^t \mathcal{B}^\eta f^\eta(u_s^\varepsilon, r_s^\varepsilon) - \mathcal{B} f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s + \int_0^t \mathcal{B} f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s.
\end{aligned}$$

Note that for the first term of the above some, a straightforward calculation gives the estimate

$$\left|\int_0^t \mathcal{B}^\eta f^\eta(u_s^\varepsilon, r_s^\varepsilon) - \mathcal{B} f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s\right| \leq 2N|E|C\mathrm{Lip}_\alpha \int_0^t \max_{i\in\mathcal{N}} |u_s^\varepsilon(\frac{i}{N}) - (u_s^\varepsilon, \psi_i^\eta)_{L^2(I)}|\mathrm{d}s$$

where $C = \max_{s\in[0,T],\varepsilon\in]0,1]} |f^\eta(u_s^\varepsilon, r_s^\varepsilon)|$ is bounded independently of η and ε according to Proposition 3.3.1 and Lip_α is a common Lipschitz constant for the $\alpha_{\xi\zeta}$'s. Using the property of the mollifiers ψ_i^η, the uniform boundedness of the family $\{u^\varepsilon, \varepsilon \in]0,1]\}$ in V and the convergence in law of u^η towards u in $\mathcal{C}([0,T], V)$, it is not difficult to see that:

$$\lim_{\eta\to 0}\lim_{\varepsilon\to 0} \mathbb{E}\left(\left|\int_0^t \mathcal{B}^\eta f^\eta(u_s^\varepsilon, r_s^\varepsilon) - \mathcal{B} f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s\right|\right) = 0.$$

It remains to show that

$$\lim_{\eta\to 0}\lim_{\varepsilon\to 0} \mathbb{E}\left|\int_0^t \mathcal{B} f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s\right| = 0.$$

We will more precisely show that:

$$\lim_{\eta\to 0}\lim_{\varepsilon\to 0} \mathbb{E}\left(\int_0^t \mathcal{B} f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s\right)^2 = 0.$$

STEP 5 (Bound the martingale part). Our previous result implies that $f^\eta \in \mathcal{D}(\mathcal{A})$ and:

$$M_t^{\eta,\varepsilon} = f^\eta(u_t^\varepsilon, r_t^\varepsilon) - f^\eta(u_0, r_0) - \int_0^t \mathcal{A} f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s \tag{3.19}$$

defines a square-integrable martingale, see for example Ethier and Kurtz [EK86], chapter 4 proposition 1.7.

Proposition 3.3.2. *We have:*

$$< M^{\eta,\varepsilon} >_t = \frac{1}{\varepsilon}\int_0^t \sum_{i\in\mathcal{N}}\sum_{\zeta\neq r_s(i)} [f^\eta(u_s^\varepsilon, r_s(r_s(i)\to\zeta)) - f^\eta(u_s^\varepsilon, r_s^\varepsilon)]^2 \alpha_{r_s(i),\zeta}(u_s^\varepsilon(\frac{i}{N}))\mathrm{d}s. \tag{3.20}$$

Proof. It is a classical proof in stochastic calculus (see for example [EK86], Chapter 1, Problem 29). The Itô and the Dynkin formulas give two distinct decompositions of the squared process $(f^\eta)^2(u_t^\varepsilon, r_t^\varepsilon)_{t\geq 0}$ in a semi-martingale. The uniqueness of the Doob-Meyer decomposition of a semi-martingale enables us to identify the bracket of our martingale. Nevertheless, we detail the proof here. Let us temporarily emphasize the role of $f^\eta \in \mathcal{D}(\mathcal{A})$ in the definition of $M^{\eta,\varepsilon}$ by writing M^{ε,f^η}. On one hand, the equation (3.19) gives for $(f^\eta)^2$:

$$(f^\eta)^2(u_t^\varepsilon, r_t^\varepsilon) = (f^\eta)^2(u_0, r_0) + \int_0^t \mathcal{A}(f^\eta)^2(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s + M_t^{\varepsilon,(f^\eta)^2} \tag{3.21}$$

On the other hand, the Itô formula (see [JS87], Theorem 4.57) applied to the process $((f^\eta)(u_t^\varepsilon, r_t^\varepsilon), t \in [0,T])$ gives:

$$\begin{aligned}(f^\eta)^2(u_t^\varepsilon, r_t^\varepsilon) &= (f^\eta)^2(u_0, r_0) + 2\int_0^t f^\eta(u_s^\varepsilon, r_{s_-}^\varepsilon)\mathrm{d}f^\eta(u_s^\varepsilon, r_s^\varepsilon) \\ &+ < f^\eta(u_\cdot^\varepsilon, r_\cdot^\varepsilon) >_t + \sum_{s\le t}[f^\eta(u_s^\varepsilon, r_s^\varepsilon) - f^\eta(u_s^\varepsilon, r_{s_-}^\varepsilon)]^2.\end{aligned}$$

Thanks to the expression (3.19) the Itô formula leads to the following semi-martingale decomposition:

$$(f^\eta)^2(u_t^\varepsilon, r_t^\varepsilon) \tag{3.22}$$

$$= (f^\eta)^2(u_0, r_0) + 2\int_0^t f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathcal{A}f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s + < M^{\varepsilon,f^\eta} >_t \tag{3.23}$$

$$+ 2\int_0^t f^\eta(u_s^\varepsilon, r_{s_-}^\varepsilon)\mathrm{d}M_s^{\varepsilon,f^\eta} + \sum_{s\le t}[f^\eta(u_s^\varepsilon, r_s^\varepsilon) - f^\eta(u_s^\varepsilon, r_{s_-}^\varepsilon)]^2. \tag{3.24}$$

We can thus identify the martingale and finite variation parts of the two expressions (3.21) and (3.22) of $(f^\eta)^2(u_t^\varepsilon, r_t^\varepsilon)$ to obtain:

$$2\int_0^t f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathcal{A}f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s + < M^{\varepsilon,f^\eta} >_t = \int_0^t \mathcal{A}(f^\eta)^2(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s,$$

providing the desired expression for the bracket. □

We have the following semi-martingale decomposition:

$$\int_0^t \mathcal{B}f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s = \varepsilon f^\eta(u_t^\varepsilon, r_t^\varepsilon) - \varepsilon f^\eta(u_0, r_0) - \varepsilon\int_0^t \frac{\mathrm{d}}{\mathrm{d}s}f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s - \varepsilon M_t^{\eta,\varepsilon}. \tag{3.25}$$

Recall that we are interested in bounding $\mathbb{E}\left|\int_0^t \mathcal{B}f^\eta(u_s^\varepsilon, r_s^\varepsilon)\mathrm{d}s\right|^2$. Since $\sup_{s\in[0,T]}|f^\eta(u_s^\varepsilon, r_s^\varepsilon)|$ is bounded independently of $\varepsilon \in]0,1]$ and $\eta > 0$, denoting by εg one of the two first terms of (3.25) we have $\mathbb{E}|\varepsilon g|^2 \le C'\varepsilon^2$ where C' is a constant independent of $\varepsilon \in]0,1]$ and $\eta > 0$. For the martingale term, by the Itô isometry:

$$\mathbb{E}(|M_t^{\eta,\varepsilon}|^2) = \mathbb{E}(< M^{\eta,\varepsilon} >_t) \le \frac{1}{\varepsilon}C',$$

since $\varepsilon < M^{\eta,\varepsilon} >_t$ is bounded uniformly in $t \in [0,T]$ and $\varepsilon \in]0,1]$ thanks to the bounds on f^η and each $\alpha_{\xi,\zeta}$. More precisely we have:

$$\begin{aligned} & < M^{\eta,\varepsilon} >_t \\ = \; & \frac{1}{\varepsilon}\int_0^t \sum_{i\in\mathcal{N}}\sum_{\zeta\neq r_s(i)} [f^\eta(u_s^\varepsilon, r_s(r_s(i)\to\zeta)) - f^\eta(u_s^\varepsilon, r_s^\varepsilon)]^2 \alpha_{r_s(i),\zeta}(u_s^\varepsilon(\frac{i}{N}))\mathrm{d}s \\ \leq \; & \frac{4T}{\varepsilon}\alpha^+ N|E| \sup_{s\in[0,T]} |f^\eta(u_s^\varepsilon, r_s^\varepsilon)|^2. \end{aligned}$$

STEP 6 (Bound the finite variation part). It remains to bound the third term of our semi-martingale decomposition.

Proposition 3.3.3. *There exists a constant C^η independent of $\varepsilon \in]0,1]$ such that:*

$$\int_0^T |\frac{\mathrm{d}}{\mathrm{d}t} f^\eta(u_t^\varepsilon, r_t^\varepsilon)|\mathrm{d}t \leq C^\eta,$$

almost surely.

Proof. Recall first in Theorem 3.2.1 the meaning of the notation $\frac{\mathrm{d}}{\mathrm{d}t}f^\eta(u_t^\varepsilon, r_t^\varepsilon)$. Recall that the map $u \mapsto f^\eta(u,r)$ for $r \in \mathcal{R}$ fixed, is Fréchet differentiable on $L^2(I)$ with Fréchet derivative in u denoted by $\frac{\mathrm{d}f^\eta}{\mathrm{d}u}(u,r)$ which is a bounded linear form on $L^2(I)$. By the Riesz representation theorem there exists $f_u^\eta(u,r) \in L^2(I)$ such that:

$$\frac{\mathrm{d}f^\eta}{\mathrm{d}u}(u,r)[h] = (f_u^\eta(u,r), h)_{L^2(I)}, \quad \forall h \in L^2(I).$$

Moreover the correspondence is isometric: $\left\|\frac{\mathrm{d}f}{\mathrm{d}u}(u,r)\right\|_{(L^2(I))^*} = \|f_u(u,r)\|_{L^2(I)}$. We know that $f_u(u,r)$ is also in V for $u \in V$ and $r \in \mathcal{R}$ and is bounded in V for bounded arguments (see Lemma 3.3.1 and Proposition 3.3.1). According to Theorem 4.iii) of [BR11] we have:

$$\frac{\mathrm{d}}{\mathrm{d}t} f^\eta(u_t^\varepsilon, r_t^\varepsilon) =< \Delta u_t^\varepsilon + G_{r_t^\varepsilon}(u_t^\varepsilon), f_u^\eta(u_t^\varepsilon, r_t^\varepsilon) > .$$

By Proposition 3.3.1, there exists a constant C_1^η independent of $\varepsilon \in]0,1]$ and $t \in [0,T]$ such that:

$$\|f_u(u_t^\varepsilon, r_t^\varepsilon)\|_V = \left\|\frac{\mathrm{d}f}{\mathrm{d}u}(u_t^\varepsilon, r_t^\varepsilon)\right\|_{V^*} \leq C_1^\eta.$$

Therefore:

$$|\frac{\mathrm{d}}{\mathrm{d}t} f(u_t^\varepsilon, r_t^\varepsilon)| \leq C_1^\eta \|\Delta u_t^\varepsilon + G_{r_t^\varepsilon}(u_t^\varepsilon)\|_{V^*}.$$

It remains to show that $\|\Delta u^\varepsilon_t + G_{r^\varepsilon_t}(u^\varepsilon_t)\|_{V^*}$ is bounded uniformly in ε and $t \in [0,T]$. For $\phi \in V$ we have, denoting by D the derivative with respect to x:

$$\begin{aligned}
& |< \Delta u^\varepsilon_t + G_{r^\varepsilon_t}(u^\varepsilon_t), \phi > | \\
= \quad & \left| - < Du^\varepsilon_t, D\phi > + \frac{1}{N} \sum_{\xi \in E} \sum_{i \in \mathcal{N}} c_\xi 1_\xi(r(i))(v_\xi - u^\varepsilon_t(\frac{i}{N}))\phi\left(\frac{i}{N}\right) \right| \\
\leq \quad & |(Du^\varepsilon_t, D\phi)_{L^2(I)}| + |E| \max_{\xi \in E} c_\xi (\max_{\xi \in E} |v_\xi| + C) C_P \|\phi\|_V \\
\leq \quad & \|Du^\varepsilon_t\|_{L^2(I)} \|D\phi\|_{L^2(I)} + |E| \max_{\xi \in E} c_\xi (\max_{\xi \in E} |v_\xi| + C) C_P \|\phi\|_V \\
\leq \quad & \|u^\varepsilon_t\|_V \|\phi\|_V + |E| \max_{\xi \in E} c_\xi (\max_{\xi \in E} |v_\xi| + C) C_P \|\phi\|_V \\
\leq \quad & (C + |E| \max_{\xi \in E} c_\xi (\max_{\xi \in E} |v_\xi| + C) C_P) \|\phi\|_V,
\end{aligned}$$

where C is the deterministic constant given by Proposition 3.2.2. Thus:

$$\|\Delta u^\varepsilon_t + G_{r^\varepsilon_t}(u^\varepsilon_t)\|_{V^*} \leq C + |E| \max_{\xi \in E} c_\xi (\max_{\xi \in E} |v_\xi| + C) C_P.$$

This ends the proof. □

STEP 7 (All bounds together). Assembling all the bounds of the different terms we see that:

$$\mathbb{E} \left| \int_0^t \mathcal{B} f^\eta(u^\varepsilon_s, r^\varepsilon_s) \mathrm{d}s \right|^2 \leq C^\eta \varepsilon$$

with the constant C^η independent of $\varepsilon \in]0,1]$ (but dependent on η). It remains to let ε goes to 0 and then η goes to 0 to conclude the proof of the identification of the limit.

STEP 8 (Uniqueness). We consider u^1 and u^2 two possible accumulation points verifying, for $i = 1, 2$:

$$u^i_t = e^{\Delta t} u_0 + \int_0^t e^{\Delta(t-s)} F(u^i_s) \mathrm{d}s$$

for all $t \in [0,T]$ almost surely. We can show easily that for any $t \in [0,T]$:

$$\mathbb{E}(\sup_{s \in [0,t]} \|u^1_s - u^2_s\|_V) \leq C \int_0^t \mathbb{E}(\sup_{l \in [0,s]} \|u^1_l - u^2_l\|_V) \mathrm{d}l$$

with the constant C independent of $t \in [0,T]$. Therefore, by application of the Gronwall's lemma we obtain that $\mathbb{E}(\sup_{t \in [0,T]} \|u^1_t - u^2_t\|_V) = 0$. □

3.3.2 The general case

We now consider the original case with different state classes $E_1, \cdots, E_l$. The first thing to do is to prove the tightness of the family $((u^\varepsilon, \bar{r}^\varepsilon), \varepsilon \in]0,1])$ in the space

$$\mathbb{D}([0,T], V \times \{1, \cdots, l\}^{|\mathcal{N}|}),$$

(see section 3.2.3 for the definition of $\bar{r}^\varepsilon$). We apply a comparison argument. We notice that:

$$\mathbb{P}(r^\varepsilon_{t+h}(i) = \zeta | r^\varepsilon_t(i) = \xi) = \left\{ \begin{array}{ll} \frac{1}{\varepsilon}\alpha_{\xi,\zeta}\left(u_t\left(\frac{i}{N}\right)\right)h + \mathrm{o}(h) & \text{if } \xi, \zeta \text{ are in the same } E_k \\ \alpha_{\xi,\zeta}\left(u_t\left(\frac{i}{N}\right)\right)h + \mathrm{o}(h) & \text{otherwise} \end{array} \right. .$$

For $\xi, \zeta \in E$ we set:

$$\lambda_{\xi,\zeta} = \left\{ \begin{array}{ll} \frac{1}{\varepsilon}\alpha^+ & \text{if } \xi, \zeta \text{ are in the same class } E_k \\ \alpha^+ & \text{otherwise} \end{array} \right. .$$

We denote by $r^{\varepsilon,\mathrm{Max}}$ the associated jump process with constant rates $\lambda_{\xi,\zeta}$. We have the following stochastic domination:

$$\mathbb{P}(r^\varepsilon_{t+h}(i) = \zeta | r^\varepsilon_t(i) = \xi) \leq \mathbb{P}(r^{\varepsilon,\mathrm{Max}}_{t+h}(i) = \zeta | r^{\varepsilon,\mathrm{Max}}_t(i) = \xi).$$

We construct the process $\bar{r}^{\varepsilon,\mathrm{Max}}$ as $\bar{r}^\varepsilon$ by aggregation. The sequence $(\bar{r}^{\varepsilon,\mathrm{Max}}, \varepsilon \in]0,1])$ is tight in $\mathbb{D}([0,T], \{1, \cdots, l\}^{|\mathcal{N}|})$, see for instance Theorem 7.4 of [YZ98]. Therefore, by comparison, the sequence $(\bar{r}^\varepsilon, \varepsilon \in]0,1])$ is also tight in

$$\mathbb{D}([0,T], \{1, \cdots, l\}^{|\mathcal{N}|}),$$

see for instance [JS87]. Moreover, the sequence $(u^\varepsilon, \varepsilon \in]0,1])$ is tight in $\mathbb{D}([0,T], V)$ by applying the arguments developed for the "all-fast" case. Endowing $V \times \{1, \cdots, l\}^{|\mathcal{N}|}$ with the product topology we see that the sequence $((u^\varepsilon, \bar{r}^\varepsilon), \varepsilon \in]0,1])$ is tight in $\mathbb{D}([0,T], V \times \{1, \cdots, l\}^{|\mathcal{N}|})$.

We must now deal with the identification of the limit. There are eight steps in the proof of the identification of the limit, we have to check that these eight steps generalize to the general case (jumping slow-fast case).

In STEP 1, by the same arguments, we obtain that it is sufficient to show that:

$$\lim_{\varepsilon \to 0} \mathbb{E}\left| \int_0^t < G_{r^\varepsilon_s}(u^\varepsilon_s) - F_{\bar{r}^\varepsilon_s}(u^\varepsilon_s), \phi > \mathrm{d}s \right| = 0.$$

Excepting notations, the arguments in STEP 2 and STEP 3 are unchanged.

In STEP 4, the Poisson equation becomes:

$$\mathcal{B}^\eta f^\eta(u,r) \quad = \quad (G^\eta_r(u) - F^\eta_{\bar{r}}(u), \phi)_{L^2(I)}, \tag{3.26}$$

where $\mathcal{B}^\eta$ is now the "fast" part of the mollified generator (f^η does not depend on $\bar{r}$ explicitly because $\bar{r}$ is constructed from r). A configuration r for the ion channels is now:

$$r \in E_{j_1} \times \cdots \times E_{j_{N-1}},$$

where $(j_1, \cdots, j_{N-1}) \in \{1, \cdots, l\}^{N-1}$ (noting that $|\mathcal{N}| = N-1$). That is each channel is in one of the class E_j for $j \in \{1, \cdots, l\}$. Then, since for fixed u the quasi-stationary measure associated to the class E_{j_k} for $k \in \{1, \cdots, N-1\}$ is $\mu_{j_k}(u)$, we have that the kernel of $(\mathcal{B}^\eta)^*$ is spanned by:

$$\{\mu^\eta_{j_1}(u) \times \cdots \times \mu^\eta_{j_{N-1}}(u), (j_1, \cdots, j_{N-1}) \in \{1, \cdots, l\}^{N-1}\}$$

when u is held fixed. Then, by the Fredholm alternative and the definition of the averaged function $F^\eta_{\bar{r}}(u)$, the Poisson equation (3.26) has a solution. Uniqueness follows by the projection condition:

$$\int_{\mathcal{R}} f^\eta(u,r)\mu^\eta_{j_1}(u)(\mathrm{d}r) \times \cdots \times \mu^\eta_{j_{N-1}}(u)(\mathrm{d}r) = 0, \quad \forall (j_1, \cdots, j_{N-1}) \in \{1, \cdots, l\}^{N-1}. \tag{3.27}$$

The proof of the last steps is then exactly the same as in the "all fast" case.

3.4 Example

In this section we give a concrete example where our result allows to reduce the complexity of a neuronal model staying nevertheless at a stochastic level.

We consider a usual Hodgkin-Huxley model but, to be very straightforward in the application of our result, we only consider the sodium current i.e. all the ion channels are sodium channels. It is still a case of interest since sodium channels are involved in the increasing phase of an action potential. At a fixed potential, the kinetic of a sodium channel is described by Figure 3.3 which represents the states and the jump rates of a continuous time Markov chain denoted by $r^\varepsilon(i)$ for the channel at position $\frac{i}{N}$ for $i \in \mathcal{N}$. A sodium channel can be in 8 different states denoted by m_ih_j for $i \in \{0,1,2,3\}$ and $j \in \{0,1\}$, the state m_3h_1 is the only "open" state for the sodium channel (see [Hil84] for more details). We simply write a_m, b_m and a_h, b_h for $a_m(u), b_m(u)$ and $a_h(u), b_h(u)$ (see Appendix 3.A for more details on the rate functions). m is the fast variable and h the slow variable for the kinetic of a sodium channel.

Each channel can therefore be in one of these eight states divided in two classes: $E = E_0 \sqcup E_1$ where

$$E_0 = \{m_0h_0, m_1h_0, m_2h_0, m_3h_0\} \text{ and } E_1 = \{m_0h_1, m_1h_1, m_2h_1, m_3h_1\}.$$

$$\begin{array}{ccccccc}
m_0h_0 & \underset{\varepsilon^{-1}b_m}{\overset{3\varepsilon^{-1}a_m}{\rightleftharpoons}} & m_1h_0 & \underset{2\varepsilon^{-1}b_m}{\overset{2\varepsilon^{-1}a_m}{\rightleftharpoons}} & m_2h_0 & \underset{3\varepsilon^{-1}b_m}{\overset{\varepsilon^{-1}a_m}{\rightleftharpoons}} & m_3h_0 \\
b_h \upharpoonleft\downharpoonright a_h & & b_h \upharpoonleft\downharpoonright a_h & & b_h \upharpoonleft\downharpoonright a_h & & b_h \upharpoonleft\downharpoonright a_h \\
m_0h_1 & \underset{\varepsilon^{-1}b_m}{\overset{3\varepsilon^{-1}a_m}{\rightleftharpoons}} & m_1h_1 & \underset{2\varepsilon^{-1}b_m}{\overset{2\varepsilon^{-1}a_m}{\rightleftharpoons}} & m_2h_1 & \underset{3\varepsilon^{-1}b_m}{\overset{\varepsilon^{-1}a_m}{\rightleftharpoons}} & m_3h_1
\end{array}$$

Figure 3.3: Kinetic of sodium channels.

Inside these two classes the states communicate at fast rates and transition between these two classes occurs at slow rates. A sodium channel is open if and only if it is in the state m_3h_1. We then define $c_{m_3h_1} = c_{\mathrm{Na}}$, $v_{m_3h_1} = v_{\mathrm{Na}}$ and $c_\xi = 0$, $v_\xi = 0$ when $\xi \in E \setminus \{m_3h_1\}$.

The equation (3.2) describing the evolution of the potential, for a number N of ion channels, is here given by :

$$\partial_t u^\varepsilon = K\Delta u^\varepsilon + \frac{1}{N}\sum_{i\in\mathcal{N}} c_{\mathrm{Na}} 1_{m_3h_1}(r^\varepsilon(i))(v_{\mathrm{Na}} - u(\frac{i}{N}))\delta_{\frac{i}{N}}. \tag{3.28}$$

The global variable u represents the difference of potential between the inside and the outside of the axon membrane. We see that Equation (3.28) is indeed of the form that we have studied in this paper. K is a constant related to the radius and the internal conductivity of the axon (see Appendix 3.A).

Let us compute the different generators and quasi-invariant measures associated to this model according to our description. The two "fast" generators are the same and are given for $u(\frac{i}{N})$ fixed (again, we do not make the dependence appear), by :

$$\mathcal{B}_j = \begin{pmatrix} -3a_m & 3a_m & 0 & 0 \\ b_m & -b_m - 2a_m & 2a_m & 0 \\ 0 & 2b_m & -2b_m - a_m & a_m \\ 0 & 0 & 3b_m & -3b_m \end{pmatrix}$$

with $j = 0$ or 1. We can compute the associated quasi-invariant measure $\mu_j(u(\frac{i}{N}))$ ($j = 0$ or 1), the only term of interest for us is :

$$\mu_1(u(\frac{i}{N}))(m_3h_1) = \frac{1}{\left(1 + \frac{b_m(u(\frac{i}{N}))}{a_m(u(\frac{i}{N}))}\right)^3}.$$

The reader familiar with the classical Hodgkin-Huxley model will notice that $\mu_1(u(\frac{i}{N}))(m_3h_1)$ is in fact equal to the steady-state function associated to the ODE describing the motion of the m-gates in the classical deterministic Hodgkin-Huxley description, see for example [HH52].

The asymptotic aggregated Markov chain $\bar{r}$ is valued in the two-state space $\{0,1\}$. According to Proposition 4.3.1 its generator is:

$$\begin{pmatrix} -a_h(u(\frac{i}{N})) & a_h(u(\frac{i}{N})) \\ b_h(u(\frac{i}{N})) & -b_h(u(\frac{i}{N})) \end{pmatrix}.$$

Therefore, according to Theorem 3.2.2, the reduced model is described by the following PDE coupled with the continuous-time Markov chain $\bar{r}$:

$$\partial_t u = K\Delta u + \frac{1}{N}\sum_{i\in\mathcal{N}} 1_1(\bar{r}(i)) c_{\mathrm{Na}}\mu_1(u(\frac{i}{N}))(m_3 h_1)(v_{\mathrm{Na}} - u(\frac{i}{N}))\delta_{\frac{i}{N}}. \tag{3.29}$$

We display a realization of the averaged piecewise deterministic Markov process in Figure 3.4. To perform the numerical simulation (in C) we extended Riedler's work (cf. [Rie12a]) to our particular framework. In [Rie12a], an algorithm (Algorithm A2) is proposed to simulate the trajectory of a PDMP in finite dimension. The idea is to simulate the jumping part of the PDMP and to use an accurate method to simulate the ODE constituting the deterministic part of the PDMP. The kinetic of the jumping component for PDMP in finite and infinite dimensions have the same form. Here we simulate the jumping part of our infinite dimensional PDMP following [Rie12a] and simulate the PDE between successive jumps by a deterministic scheme. For the PDE, we used an explicit finite difference Euler scheme in space and time. Simulating the jumps of an inhomogeneous time Markov chain is classical and there exist a lot of efficient numerical schemes for PDE. Therefore the described method is natural. This argument is of course not a proof but gives an heuristic interpretation of our approach.

In Figure 3.4, by $N = \infty$, we mean the limit of the averaged model (3.29) when N goes to infinity. According to [Aus08], the model (3.29) should converge in distribution when N goes to infinity towards the following deterministic model:

$$\left\{ \begin{array}{rcl} \partial_t u^{(\infty)} & = & K\Delta u^{(\infty)} + h c_{\mathrm{Na}}\mu_1(u^{(\infty)})(m_3 h_1)(v_{\mathrm{Na}} - u^{(\infty)}), \\ \partial_t h & = & (1-h)a_h(u^{(\infty)}) - h b_h(u^{(\infty)}). \end{array} \right. \tag{3.30}$$

This convergence is illustrated numerically in Figure 3.4. We notice that the mean speed γ_N of the front wave generated by model (3.29) for a fixed N is about $\gamma_\infty - \frac{\alpha}{\sqrt{N}}$ where γ_∞ is the mean speed of the front wave for the model (3.30) and α is a positive real. It means that the renormalized difference $\sqrt{N}(\gamma_\infty - \gamma_N)$ is of order one. This is consistent with recent results of [RT13] where, for model of type (3.29) but in replacing the Dirac masses by functions at least in $L^2(I)$, the authors show that the renormalized difference $\sqrt{N}(u^{(N)} - u^{(\infty)})$ converges in distribution.

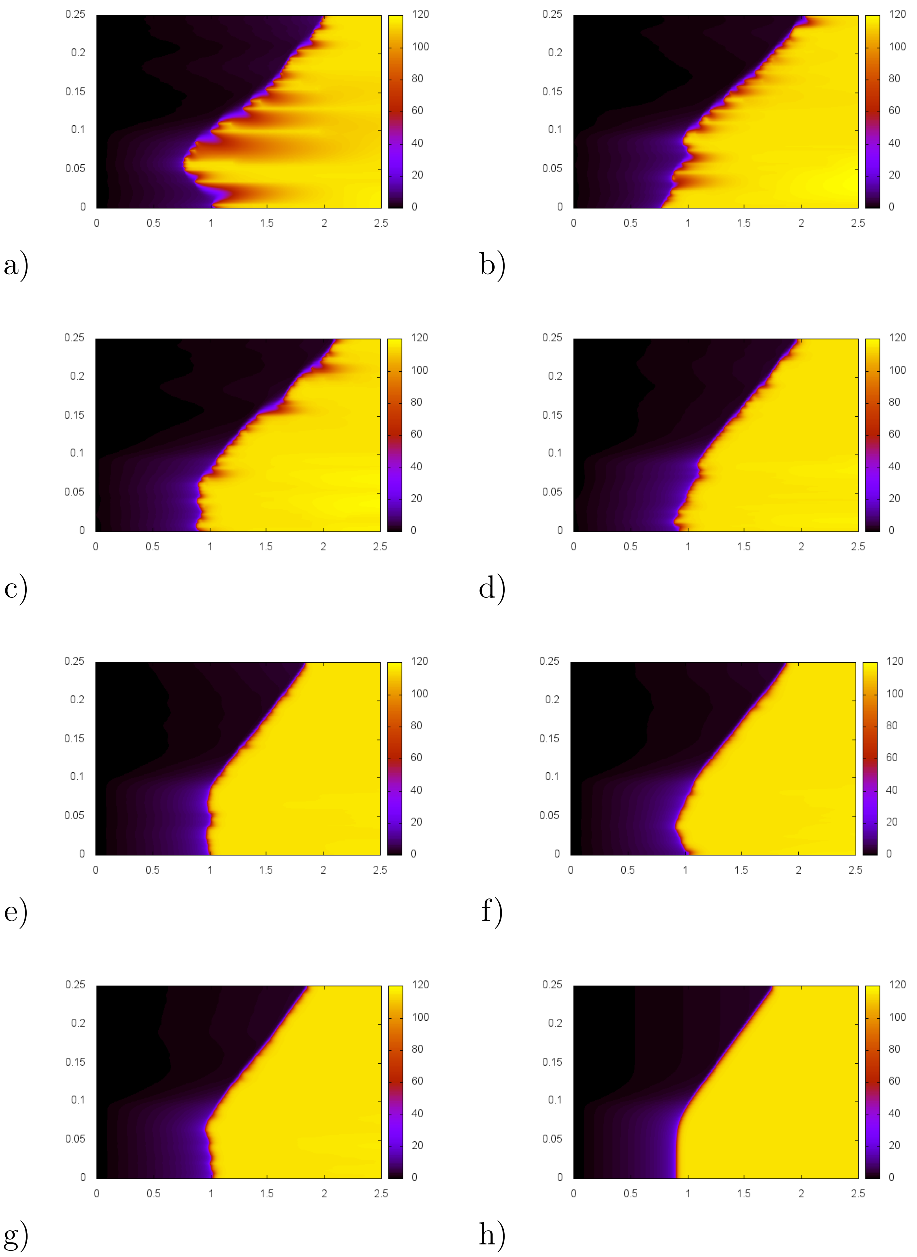


Figure 3.4: Simulation of the action potential against space (vertical axis) and time (horizontal axis). The averaged model (3.29) is displayed for various number of ion channels: a) $N = 50$, b) $N = 100$, c) $N = 150$, d) $N = 250$, e) $N = 500$, f) $N = 700$, g) $N = 900$, h) $N = \infty$. We see what we expect from the partial differential equation: a traveling wave connecting the initial state 0 to the stable state v_{Na}.

Appendix 3.A Functions and data for the simulation

For the simulation, the jump rate functions are given, for real u, by :

$$\begin{aligned}
a_m(u) &= \frac{0.1(25-u)}{e^{2.5-0.1u}-1}\\
b_m(u) &= 4e^{\frac{-u}{18}}\\
a_h(u) &= 0.07e^{\frac{-u}{20}}\\
b_h(u) &= \frac{1}{e^{3-0.1u}+1}.
\end{aligned}$$

The maximal conductance associated to sodium ions is $c_{\mathrm{Na}} = 120$ mS.cm^{-2} and the potential at rest is $v_{\mathrm{Na}} = 115$ mV. The constant K is given by $K = \frac{a}{2R}$ where a is the radius of the axon, $a = 0.0238$ cm and R is the internal resistance of the axon, $R = 34.5$ Ω.cm, these data are classical, see for example [HH52]. We have used an input on the potential equal to $\mu = 6.7$ all along the time on the segment $[0, 0.1]$ of the axon.

Chapter 4

Asymptotic normality for a class of Hilbert-valued piecewise deterministic Markov processes

The material for chapter 4 is taken from the submitted pre-publication [GT13a] *Multiscale Piecewise Deterministic Markov Process in Infinite Dimension: Central Limit Theorem* available on Arxiv. Notations have been modified to match up with the other chapters of the present text. Since the present thesis is provided with a general introduction in Chapter 1 and especially for this chapter in Section 1.3.2, we start with a shorter introduction than in the corresponding pre-publication.

4.1 Introduction

In Chapter 3, we addressed the question of averaging for a class of multiscale spatially extended stochastic conductance-based neuron models, also known as spatially extended stochastic generalized Hodgkin-Huxley models. These models describe the evolution of an action potential or nerve impulse along the axon of a neuron at the scale of ionic channels. More generally, in electro-physiology, these equations describe the evolution of an action potential in excitable membranes. Mathematically, these spatially extended stochastic conductance-based models belong to the class of Hilbert-valued Piecewise Deterministic Markov Processes (PDMP) with multiple time scales. We obtained averaging results for this class of models. The averaged models are still Hilbert-valued PDMPs but of lower dimensions in the sense that the dynamic of the jump components of the slow-fast PDMP is simplified. In the present chapter, we study the fluctuations of the original slow-fast systems around their averaged limit. A central limit theorem is derived. A numerical example based on a spatially extended stochastic

Morris-Lecar model is provided at the end of the chapter.

The chapter is organized as follows. In Section 4.2 and 4.3 we present the model and recall as briefly as possible the main results of Chapter 3 and in particular the different properties of the averaged process. Section 4.4 introduces the main result of the present chapter: the Central Limit Theorem. The description of the general class of PDMP which can be included in our framework is described in Section 4.2.3. In Section 4.5, we begin by proving the Central Limit Theorem in the so-called all-fast case of Section 4.4.1. In the all fast case, we divide the proof in two parts: tightness in Section 4.5.1 and identification of the limit in Section 4.5.2. Properties of the diffusion operator related to the fluctuations are investigated in Section 4.5.3. In Section 4.6, as an example, we consider a spatially extended stochastic Morris-Lecar model and provide numerical experiments.

4.2 The models

4.2.1 Stochastic Hodgkin-Huxley models

In this section, we introduce the stochastic generalized Hodgkin-Huxley model also known in the literature as stochastic conductance-based neuron models. This model was first considered in [Aus08], and later in [BR11, GT12, RTW12]. Although we are interested in multiscale stochastic conductance-based neuron models, we start by describing the model without different time scales, for the sake of clarity. We begin by stating all our mathematical definitions and assumptions before providing the biological interpretation of our model.

Let T be a fixed finite time horizon, $I = [0, 1]$ and E a finite set. We fix an integer $N \geq 1$ and consider the subset $\mathcal{N} = \{z_i, i = 1, 2, \ldots, N\}$ of $\mathring{I} = (0, 1)$. We write $\mathcal{R}$ for the finite set $E^{\mathcal{N}}$ and $V = H_0^1(I)$ for the space of functions in $L^2(I)$ with first distributional derivative also belonging to $L^2(I)$. Remember that the Hilbert spaces V and $L^2(I)$, the Laplacian operator Δ and the Dirac delta function on V are defined in Chapter 2, Section 2.1. Let us simply recall that V and $L^2(I)$ are both Hilbert spaces with respective scalar products denoted by $(\cdot, \cdot)$ and $(\cdot, \cdot)_{L^2(I)}$.

For $(u, r) \in V \times \mathcal{R}$ we define the generalized function $G_r(u)$ (or reaction term) in V^* by

$$G_r(u) = \frac{1}{N} \sum_{i=1}^{N} c_{r(i)}(v_{r(i)} - u(z_i))\delta_{z_i}, \tag{4.1}$$

where $V^* = H^{-1}(I)$ is the dual space of V. We denote by $< \cdot, \cdot >$ the duality pairing between V and V^*. For $\xi \in E$, c_ξ and v_ξ are two real constants, the first being positive. We omit N in the notation of $G_r(u)$ because, in contrary to [Aus08, RTW12], N is held fixed all along the chapter. Notice that, in contrary to the model developed in Section 4.2.2, $G_r(u)$ does not belong to V, thus, the model of the present section does not enter in the general framework of Section 4.2.3. However, we prefer to present in a first part the model given by (4.1) with the Dirac delta functions because it corresponds exactly to the model studied in [Aus08, GT12].

For two states $\xi, \zeta \in E$, we define by $\alpha_{\xi,\zeta}$ the jump intensity or transition rate function from the state ξ to the state ζ. The function $\alpha_{\xi,\zeta}$ is a real valued function of a real variable supposed to be, as its derivative, Lipschitz-continuous. We assume moreover that $0 \leq \alpha_{\xi,\zeta} \leq \alpha^+$ for any $\xi, \zeta \in E$ and either $\alpha_{\xi,\zeta}$ is constant equal to zero or is positive bounded below by a positive constant α_-. That is, the non-zero rate functions are bounded below and above by positive constants.
Then, for $u \in V$ and $(r, \tilde{r})$ two different states of $\mathcal{R}$, we define by $q_{r\tilde{r}}$ the jump intensity or transition rate function from the state r to the state $\tilde{r}$. This is a real valued function defined on V by

$$q_{r\tilde{r}}(u) = \begin{cases} 0 & \text{if } r \text{ and } \tilde{r} \text{ differ from more than one component,} \\ \dfrac{\alpha_{r(i)\tilde{r}(i)}(u(z_i))}{\alpha_{r(i)}(u(z(i)))} & \text{if } r(i) \neq \tilde{r}(i) \text{ and all the other components are equal.} \end{cases} \tag{4.2}$$

The quantity $\alpha_{r(i)}(u(z_i)) = \sum_{\xi \in E \setminus \{r(i)\}} \alpha_{r(i)\xi}(u(z_i))$ represents the total rate of leaving the state $r(i) \in E$.
The stochastic conductance-based model for excitable cells we consider consists in the following evolution problem on I

$$\begin{cases} \partial_t u_t &= \Delta u_t + G_{r_t}(u_t), \\ \mathbb{P}(r_{t+h} = \tilde{r} | r_t = r) &= q_{r\tilde{r}}(u_t)h + \mathrm{o}(h) \end{cases} \tag{4.3}$$

for $t \in [0, T]$ and zero Dirichlet boundary conditions. That is $u_t(0) = u_t(1) = 0$ for all $t \in [0, T]$. We are interested in the stochastic process $(u_t, r_t)_{t \in [0,T]}$.

The spatially extended stochastic Hodgkin-Huxley model (4.3) describes the propagation of an action potential along an axon at the scale of ionic channels. The axon, or nerve fiber, is the component of a neuron which allows the propagation of an incoming signal from the soma to another neuron on long distances. The length of the axon is large relative to its radius, thus, for mathematical convenience, we consider the axon as a segment I. All along the axon are the ion channels which allow and amplify the propagation of the incoming impulse. We assume that there are N ion channels along the axon located in the subset $\mathcal{N} = \{z_i, i = 1, 2, \ldots, N\}$ of $\mathring{I} = (0, 1)$. In [Aus08, GT12] for instance, $\mathcal{N} = \{\frac{i}{N}, i = 1, \ldots, N-1\}$ which means that the ion channels are regularly spaced. Each ion channel can be in several states $\xi \in E$, for instance, in the Hodgkin-Huxley model, a state can be: "receptive to sodium ions and open". When a ion channel is open, it allows some ionic species to enter or leave the cell, generating in this way a current. For a greater insight into the underlying biological phenomena governing the model, the authors refer to [Hil84], Chapter 2.

The ion channels switch between states according to a continuous time Markov chain whose jump intensities depend on the local potential of the axon membrane. For a given channel, the rate function describes the rate at which it switches from one state to another.
A possible configuration of all the N ion channels is denoted by $r = (r(i), i \in \mathcal{N})$, a point in the space of all configurations $\mathcal{R} = E^{\mathcal{N}}$: $r(i)$ is the state of the channel located at z_i, for $i \in \mathcal{N}$. The channels, or stochastic processes $r(i)$, are supposed to evolve independently over infinitesimal timescales. Denoting by $u_t(z_i)$ the local potential at point z_i at time t, we have

$$\mathbb{P}(r_{t+h}(i) = \zeta | r_t(i) = \xi) = \alpha_{\xi,\zeta}\left(u_t\left(z_i\right)\right) h + \mathrm{o}(h). \tag{4.4}$$

For any $\xi \in E$, c_ξ represents the maximal conductance and v_ξ the steady state potentials, or driven potentials, of a channel in state ξ.
The transmembrane potential $u_t(x)$, that is the difference of electrical potential between the outside and the inside of the axon, evolves according to the following hybrid reaction-diffusion PDE

$$\partial_t u_t = \Delta u_t + \frac{1}{N}\sum_{i=1}^{N} c_{r_t(i)}(v_{r_t(i)} - u_t(z_i))\delta_{z_i}. \tag{4.5}$$

The zero Dirichlet boundary conditions for this PDE corresponds to the case of a clamped axon [Hil84].

4.2.2 Stochastic Hodgkin-Huxley models with mollifiers

For technical reasons, in the present chapter, we will work with a slightly different model where the Dirac distributions δ_{z_i} in (4.5) are replaced by approximations ϕ_{z_i} in the sense of distributions, in the same way as in so called compartment models. In such a model the reaction term is given by

$$G_r(u) = \frac{1}{N}\sum_{i=1}^{N} c_{r(i)}(v_{r(i)} - \bar{u}_i)\phi_{z_i} \tag{4.6}$$

for $(r, u) \in \mathcal{R} \times L^2(I)$ and where, for any $h \in L^2(I)$: $\bar{h}_i = (h, \phi_{z_i})_{L^2(I)}$. For $i \in \mathcal{N}$, the function ϕ_{z_i} which belongs to $L^2(I)$ approximates the Dirac distribution δ_{z_i}.

For $i \in \{1, \dots, N\}$ the functions ϕ_{z_i} are defined on I by

$$\phi_{z_i}(x) = \frac{1}{\kappa} M\left(\frac{x - z_i}{\kappa}\right)$$

with κ small enough such that ϕ_{z_i} is compactly supported in I. The mollifier M is defined on $\mathbb{R}$ by

$$M(x) = e^{-\frac{1}{1-x^2}} 1_{[-1,1]}(x).$$

Replacing δ_{z_i} by ϕ_{z_i} corresponds to consider that when the channel located at z_i is open and allows a current to pass, not only the voltage at the point z_i is affected, but also the voltage on a small area around z_i, see [BR11], Section 3.1. The family of functions ϕ_{z_i} is indexed by a parameter κ related to the considered membrane area: the smaller κ is, the smaller is the area. When u is held fixed, the dynamic of the ion channel at location z_i is given by

$$\mathbb{P}(r_{t+h}(i) = \zeta | r_t(i) = \xi) = \alpha_{\xi,\zeta}(\bar{u}_i)\, h + \mathrm{o}(h) \tag{4.7}$$

for $\xi, \zeta \in E$ and $t, h \geq 0$.

The present paper is thus concerned with the following evolution problem, for $t \in [0, T]$

$$\left\{ \begin{array}{rcl} \partial_t u_t & = & \Delta u_t + G_{r_t}(u_t), \\ \mathbb{P}(r_{t+h}(i) = \zeta | r_t(i) = \xi) & = & \alpha_{\xi,\zeta}(\bar{u}_i)\, h + \mathrm{o}(h). \end{array} \right. \tag{4.8}$$

4.2.3 A general framework

The previous stochastic Hodgkin-Huxley model with mollifiers actually belongs to a more general framework that we now describe.
Let A be a self-adjoint linear operator on a separable Hilbert space H with associated norm $\|\cdot\|_H$ such that there exists a Hilbert basis $\{e_k, k \geq 1\}$ of H made up with eigenvectors of A

$$Ae_k = -l_k e_k \tag{4.9}$$

for $k \geq 1$ and such that

$$\sup_{k \geq 1} \sup_{y \in I} |e_k(y)| < \infty. \tag{4.10}$$

The eigenvalues $\{l_k, k \geq 1\}$ are assumed to form an increasing sequence of positive numbers enjoying the following property

$$\sum \frac{1}{l_k} < \infty. \tag{4.11}$$

Let $\mathcal{R}$ be a finite set. For any $r \in \mathcal{R}$, the reaction term $G_r : H \mapsto H$ is globally Lipschitz on H uniformly on $r \in \mathcal{R}$. That is to say, there exists a constant $L_G > 0$ such that for any $(r, u, \tilde{u}) \in \mathcal{R} \times H \times H$ we have

$$\|G_r(u) - G_r(\tilde{u})\|_H \leq L_G \|u - \tilde{u}\|_H. \tag{4.12}$$

For fixed $u \in H$ let $Q(u) := (q_{r\tilde{r}}(u))_{(r,\tilde{r}) \in \mathcal{R}\times\mathcal{R}}$ be the generator of a continuous time Markov chain $(r_t, t \geq 0)$ on $\mathcal{R}$. We assume that for $r \neq \tilde{r}$, the intensity rate functions $q_{r\tilde{r}} : H \mapsto \mathbb{R}_+$ are uniformly bounded and Lipschitz continuous. There exist two constants B_q, L_q such that for any $(r, \tilde{r}, u, \tilde{u}) \in \mathcal{R} \times \mathcal{R} \times H \times H$ we have

$$\sup_{(r,\tilde{r})\in\mathcal{R}\times\mathcal{R}} \sup_{u \in H} q_{r\tilde{r}}(u) \leq B_q, \quad |q_{r\tilde{r}}(u) - q_{r\tilde{r}}(\tilde{u})| \leq L_q \|u - \tilde{u}\|_H. \tag{4.13}$$

Moreover, we assume that there exists a positive constant q_- such that

$$\inf_{u \in H} \lambda(u) \geq q_-, \tag{4.14}$$

where $\lambda(u)$ is the first non-zero eigenvalue of $Q(u)$. We also assume that there exists a unique pseudo-invariant measure $\mu(u)$ associated to the generator $Q(u)$ which is bounded and Lipschitz continuous with respect to u.
The present paper is concerned with the following evolution problem, for $t \in [0, T]$

$$\left\{ \begin{array}{rcl} \partial_t u_t & = & A u_t + G_{r_t}(u_t), \\ \mathbb{P}(r_{t+h} = \tilde{r} | r_t = r) & = & q_{r\tilde{r}}(u_t) h + o(h). \end{array} \right. \tag{4.15}$$

Let us mention that in this framework, the model with mollifiers corresponds to $H = L^2(I)$ and $\mathcal{R} = E^N$. With $A = \Delta$, the Hilbert space basis $\{f_k, k \geq 1\}$ of $L^2(I)$ defined in Chapter 2, Section 2.1 and $l_k = (k\pi)^2$ for $k \geq 1$, Assumptions (4.9)-(4.14) are satisfied.

4.2.4 Basic properties of stochastic Hodgkin-Huxley models

The following result states that there exists a stochastic process satisfying system (4.8). Let u_0 be in $\mathcal{D}(\Delta)$ such that $\min_{\xi \in E} v_\xi \leq u_0 \leq \max_{\xi \in E} v_\xi$, the initial potential of the axon. Let $q_0 \in \mathcal{R}$ be the initial configuration of the ion channels.

Proposition 4.2.1 ([BR11]). *Fix $N \geq 1$. There exists a pair $(u_t, r_t)_{0 \leq t \leq T}$ of càdlàg adapted stochastic processes satisfying that each sample path of u is in $\mathcal{C}([0, T], L^2(I))$, r_t is in $\mathcal{R}$ for all $t \in [0, T]$ and $(u_t, r_t)_{0 \leq t \leq T}$ is solution of (4.8). Moreover $(u_t, r_t)_{0 \leq t \leq T}$ is a so called Piecewise Deterministic Markov Process.*

The existence of a stochastic process solution of (4.3) has been first proved in [Aus08] for the model with Dirac mass. The proof in [Aus08] is in two parts. First, the Schaeffer fixed point theorem implies that when the jump process r jumps at rate 1, there exists a solution to (4.3). Then the original dynamic of r is recovered using the Girsanov theorem for càdlàg processes with finite state space. Another approach has been developed in [BR11]. There, the process (u, r) is constructed explicitly as a piecewise deterministic Markov process generalizing in this way the theory developed by Davis [Dav84, Dav93] from the finite to the infinite dimensional setting. In [BR11], the authors prove that their process is Markovian and moreover characterize its generator. Still, another approach based on the marked point process theory is also possible, see for instance [Jac05], Chapter 7 and the extension to our framework in [Rie12b].

We proceed now by recalling the form of the generator of the process $(u_t, r_t)_{0\leq t\leq T}$ solution of (4.3). For $(u_0, r) \in L^2(I) \times \mathcal{R}$, we denote by $(\psi_r(t, u_0), t \in [0, T])$ the unique solution starting from u_0 of the PDE

$$\partial_t u_t = \Delta u_t + \frac{1}{N}\sum_{i=1}^{N} c_{r(i)}(v_{r(i)} - \overline{u_{t}}_i)\phi_{z_i} \tag{4.16}$$

with zero Dirichlet boundary conditions.

Proposition 4.2.2. *Let f be a locally bounded measurable function on $L^2(I) \times \mathcal{R}$ such that the map $t \mapsto f(\psi_r(t, u_0), r)$ is absolutely-continuous for all $(u_0, r) \in L^2(I) \times \mathcal{R}$. Then f is in the domain $\mathcal{D}(\mathcal{A})$ of the extended generator of the process (u, r). The extended generator is given for almost all t by*

$$\mathcal{A}f(u_t, r_t) = \frac{\mathrm{d}f}{\mathrm{d}t}(u_\cdot, r_t)(t) + \mathcal{B}(u_t)f(u_t, \cdot)(r_t), \tag{4.17}$$

where

$$\mathcal{B}(u_t)f(u_t, \cdot)(r_t) = \sum_{i=1}^{N}\sum_{\zeta\in E}[f(u_t, r_t(r_t(i) \to \zeta)) - f(u_t, r_t)]\alpha_{r_t(i),\zeta}(\overline{u_{t}}_i).$$

The element $r_t(r_t(i) \to \zeta)$ of $\mathcal{R}$ is equal to $r_t(j)$ if $j \neq i$ and to ζ if $j = i$. The notation $\frac{\mathrm{d}}{\mathrm{d}t}f(u_\cdot, r_t)(t)$ means that the function $s \mapsto f(u'_s, r)$ is differentiated at $s = t$, where u' is the solution of the PDE (4.16) with the channel state r_t held fixed equal to r. When f is continuously Fréchet differentiable with respect to its first argument and such that the Riesz representation $f_u \in L^2(I)$ of the Fréchet derivative satisfies $f_u(u, r) \in V$ for $u \in V$ and is a locally bounded composition operator in $L^2((0, T), V)$ then

$$\frac{\mathrm{d}f}{\mathrm{d}t}(u_\cdot, r_t)(t) = < f_u(u_t, r_t), \Delta u_t + G_{r_t}(u_t) > .$$

See Chapter 2 Section 2.4 for the definition and main properties of Fréchet differentiable functions.

4.3 Multiscale models, singular perturbation and averaging

In this section, we introduce a slow-fast dynamic in the stochastic Hodgkin-Huxley model described in Section 4.2.2: some states of the ion channels communicate faster between each other than others. This is biologically relevant as remarked for example in [Hil84], Chapter 18. Mathematically, this leads to the introduction of an additional small parameter $\varepsilon > 0$ in our previously described model: the states which communicate at a faster rate communicate at the previous rate $\alpha_{\xi,\zeta}$ divided by ε. For an introduction on slow-fast systems, we refer to [PS08], for a general theory of slow-fast continuous time Markov chain, see [YZ98] and for the case of slow-fast systems with diffusion, see [BG06].

In the context of Section 4.2.2, we make a partition of the state space E according to the different orders in ε of the rate functions

$$E = E_1 \sqcup \cdots \sqcup E_l,$$

where $l \in \{1, 2, \cdots\}$ is the number of classes. Inside a class E_j, the states communicate faster at jump rates of order $\frac{1}{\varepsilon}$. States in different classes communicate at the usual rate of order 1. For $\varepsilon > 0$ fixed, we denote by $(u^\varepsilon, r^\varepsilon)$ the modification of the PDMP introduced in the previous section with now two time scales. Its generator is, for $f \in \mathcal{D}(\mathcal{A}^\varepsilon)$

$$\mathcal{A}^\varepsilon f(u^\varepsilon_t, r^\varepsilon_t) = \frac{\mathrm{d}f}{\mathrm{d}t}(u^\varepsilon_\cdot, r^\varepsilon_t)(t) + \mathcal{B}^\varepsilon(u^\varepsilon_t) f(u^\varepsilon_t, \cdot)(r^\varepsilon_t). \tag{4.18}$$

The term $\mathcal{B}^\varepsilon$ is the component of the generator related to the continuous time Markov chain r^ε. According to (4.17) and our slow-fast description, we have the two time scales decomposition of this generator

$$\mathcal{B}^\varepsilon = \frac{1}{\varepsilon}\mathcal{B} + \hat{\mathcal{B}}, \tag{4.19}$$

where the "fast" generator $\mathcal{B}$ is given by

$$\begin{aligned}&\mathcal{B}(u^\varepsilon_t) f(u^\varepsilon_t, r^\varepsilon_t)\\ &= \sum_{i=1}^N \sum_{j=1}^l 1_{E_j}(r^\varepsilon_t(i)) \sum_{\zeta \in E_j} [f(u^\varepsilon_t, r^\varepsilon_t(r^\varepsilon_t(i) \to \zeta)) - f(u^\varepsilon_t, r^\varepsilon_t)] \alpha_{r^\varepsilon_t(i),\zeta}(\overline{u^\varepsilon_t}_i)\end{aligned} \tag{4.20}$$

and the "slow" generator $\hat{\mathcal{B}}$ is given by

$$\begin{aligned}&\hat{\mathcal{B}}(u_t^\varepsilon)f(u_t^\varepsilon, r_t^\varepsilon)\\&=\sum_{i=1}^N\sum_{j=1}^l 1_{E_j}(r_t^\varepsilon(i))\sum_{\zeta\notin E_j}[f(u_t^\varepsilon, r_t^\varepsilon(r_t^\varepsilon(i)\to\zeta))-f(u_t^\varepsilon, r_t^\varepsilon)]\alpha_{r_t^\varepsilon(i),\zeta}(\overline{u_t^\varepsilon}_i).\end{aligned} \tag{4.21}$$

For $y \in \mathbb{R}$ fixed and $g : \mathbb{R} \times E \to \mathbb{R}$, we denote by $\mathcal{B}_j(y)$, $j \in \{1, \cdots, l\}$ the following generator

$$\mathcal{B}_j(y)g(\xi) = 1_{E_j}(\xi)\sum_{\zeta\in E_j}[g(y,\zeta)-g(y,\xi)]\alpha_{\xi,\zeta}(y).$$

For any $y \in \mathbb{R}$ fixed, and any $j \in \{1, \cdots, l\}$, we assume that the fast generator $\mathcal{B}_j(y)$ is weakly irreducible on E_j, i.e. has a unique quasi-stationary distribution denoted by $\mu_j(y)$. This quasi-stationary distribution is supposed to be Lipschitz-continuous in y, as well as its derivative.

Following [YZ98], the states in E_j can be considered as equivalent. For any $i = 1, \dots, N$ we define a new stochastic process $(\bar{r}_t^\epsilon)_{t\geq 0}$ by $\bar{r}_t^\varepsilon(i) = j$ when $r_t^\varepsilon(i) \in E_j$ and abbreviate E_j by j . We then obtain an aggregate process $\bar{r}^\varepsilon(i)$ with values in $\{1, \cdots, l\}$. This process is also often called the coarse-grained process. It is not a Markov process for $\varepsilon > 0$ but a Markovian structure is recovered at the limit when ε goes to 0. More precisely, we have the following proposition.

Proposition 4.3.1 ([YZ98], Chapter 7)**.** *For any $y \in \mathbb{R}$, $i = 1, \dots, N$, the process $\bar{r}^\varepsilon(i)$ converges weakly when ε goes to 0 to a Markov process $\bar{r}(i)$ generated by*

$$\bar{\mathcal{B}}(y)g(\bar{r}(i)) = \sum_{j=1}^l 1_j(\bar{r}(i))\sum_{k=1,k\neq j}^l (g(k)-g(j))\sum_{\xi\in E_j}\sum_{\zeta\in E_k}\alpha_{\zeta,\xi}(y)\mu_j(y)(\zeta)$$

with $g : \{1, \cdots, l\} \to \mathbb{R}$ measurable and bounded.

We proved in [GT12] (in the context of the model with Dirac mass) that the limit of $(u^\varepsilon, \bar{r}^\varepsilon)$ when ε goes to zero requires to average the reaction term $G_r(u)$ against the quasi-invariant distributions. That is we consider that each cluster of states E_j has reached its stationary behavior. This leads to the averaged reaction term of the following form: for any $\bar{r} \in \bar{\mathcal{R}} = \{1, \cdots, l\}^N$

$$F_{\bar{r}}(u) = \frac{1}{N}\sum_{i=1}^N\sum_{j=1}^l 1_j(\bar{r}(i))\sum_{\zeta\in E_j} c_\zeta\mu_j(\bar{u}_i)(\zeta)(v_\zeta-\bar{u}_i)\phi_{z_i}. \tag{4.22}$$

Therefore, we call the following hybrid PDE

$$\partial_t u_t = \Delta u_t + F_{\bar{r}_t}(u_t), \tag{4.23}$$

the *averaged equation* of (4.5). We take zero Dirichlet boundary conditions and initial conditions u_0 and $\bar{q}_0$ where $\bar{q}_0$ is the aggregation of the initial channel configuration q_0: $\bar{q}_0 = \sum_{j=1}^{l} j 1_{E_j}(q_0)$. In equation (4.23), each coordinate of $(\bar{r}_t)_{t\in[0,T]}$ evolves independently over infinitesimal time intervals and according to the averaged jump rates between the subsets E_j of E. For j and k in $\{1, \cdots, l\}$, the average jump rate from class E_j to class E_k is given by

$$\bar{\alpha}_{jk}(y) = \sum_{\zeta \in E_j} \sum_{\xi \in E_k} \alpha_{\zeta,\xi}(y)\mu_j(y)(\zeta). \tag{4.24}$$

We can now state the averaging result proved for the model with Dirac mass in Chapter 3 but easily adaptable to the model with mollifiers.

Theorem 4.3.1. *When ε goes to 0 the stochastic process $(u^\varepsilon, \bar{r}^\varepsilon)$ solution of (4.18) converges in distribution in the space $\mathcal{C}([0,T], L^2(I)) \times \mathbb{D}([0,T], \mathcal{R})$ to $(u, \bar{r})$, solution of (4.23)-(4.24).*

Let us recall a result of first importance to prove Theorem 4.3.1 and in the present paper as well. We refer the interested reader to Chapter 3 for the proof. This result establishes the uniform boundedness in ε of the process u^ε.

Proposition 4.3.2. *For any $T > 0$, there is a deterministic positive constant C independent of $\varepsilon \in]0,1]$ such that*

$$\sup_{t\in[0,T]} \|u_t^\varepsilon\|_{L^2(I)} \leq C,$$

almost-surely.

For the sake of completeness, we recall a second result which states that the averaged model is well posed and is still a PDMP.

Proposition 4.3.3. *For any $T > 0$ there exists a probability space such that equations (4.23)-(4.24) define a PDMP $(u_t, \bar{r}_t)_{t\in[0,T]}$ in infinite dimension in the sense of [BR11]. Moreover, there is a constant C such that*

$$\sup_{t\in[0,T]} \|u_t\|_{L^2(I)} \leq C$$

and $u \in \mathcal{C}([0,T], L^2(I))$ almost-surely.

4.4 Main results

We present in this section the main results of the present paper. The averaging result of Theorem 4.3.1 above may be seen as a Law of Large Numbers. The natural next step is then to study the fluctuations of the slow-fast system around its averaged limit, in other words, to look for a Central Limit Theorem.

4.4.1 Fluctuations for the stochastic Hodgkin-Huxley models

For the sake of clarity in our presentation, we first present our result in the so called all-fast case that we proceed to define.
When all states in E communicate at fast rates, there is a single class as described in Section 4.3, which is equal to the whole set E. For each $\varepsilon > 0$, the generator of the process $(u^\varepsilon, r^\varepsilon)$ is given by

$$\mathcal{A}^\varepsilon f(u_t^\varepsilon, r_t^\varepsilon) = \frac{\mathrm{d}f}{\mathrm{d}t}(u_\cdot^\varepsilon, r_t^\varepsilon)(t) + \frac{1}{\varepsilon}\mathcal{B}(u_t^\varepsilon) f(u_t^\varepsilon, \cdot)(r_t^\varepsilon), \tag{4.25}$$

where the slow part of the generator reduces to zero, $\hat{\mathcal{B}} \equiv 0$ in Section 4.3, and

$$\mathcal{B}(u_t^\varepsilon) f(u_t^\varepsilon, r_t^\varepsilon) = \sum_{i=1}^N \sum_{\xi\in E} [f(u_t^\varepsilon, r_t^\varepsilon(r_t^\varepsilon(i) \to \xi)) - f(u_t^\varepsilon, r_t^\varepsilon)]\alpha_{r_t^\varepsilon(i),\xi}(\overline{u_t^\varepsilon}_i).$$

When $u \in V$ is held fixed, the Markov process $r(i)$ has a unique stationary distribution $\mu(\bar{u}_i)$ for any $i = 1, \dots, N$. Then the process $(r(i), i = 1, \dots, N)$ has the following stationary distribution

$$\mu(u) = \bigotimes_{i=1}^N \mu(\bar{u}_i).$$

The averaged reaction term reduces to

$$F(u) = \int_{\mathcal{R}} G_r(u)\mu(u)(\mathrm{d}r) = \frac{1}{N}\sum_{\xi\in E}\sum_{i=1}^N c_\xi \mu(\bar{u}_i)(\xi)(v_\xi - \bar{u}_i)\phi_{z_i}. \tag{4.26}$$

The averaged limit u is solution of the PDE

$$\partial_t u_t = \Delta u_t + F(u_t)$$

with initial condition u_0 and zero Dirichlet boundary conditions. Note that in this case, the limit PDE is no longer hybrid in contrast with (4.23). For $\varepsilon > 0$, we denote by z^ε the renormalized difference between u^ε and u:

$$z^\varepsilon = \frac{u^\varepsilon - u}{\sqrt{\varepsilon}}. \tag{4.27}$$

The main result of the present paper is the following.

Theorem 4.4.1. *When ε goes to 0 the process z^ε converges in distribution in $\mathcal{C}([0,T], L^2(I))$ towards a process z. For $u \in L^2(I)$, let $C(u) : L^2(I) \to L^2(I)$ be a diffusion operator characterized by*

$$(C(u)f_j, f_i)_{L^2(I)} = \int_{\mathcal{R}} (G_r(u) - F(u), f_i)_{L^2(I)} (\Phi(r,u), f_j)_{L^2(I)} \mu(u)(\mathrm{d}r).$$

Φ is the unique solution of the equation

$$\begin{cases} \mathcal{B}(u)\Phi(r,u) &= -(G_r(u) - F(u)) \\ \int_{\mathcal{R}} \Phi(r,u)\mu(u)(\mathrm{d}r) &= 0. \end{cases} \tag{4.28}$$

Let us also define an operator $\bar{\mathcal{G}}^1$ by, for $t \in [0,T]$ and a measurable, bounded and twice Fréchet differentiable function $\psi : L^2(I) \to \mathbb{R}$,

$$\bar{\mathcal{G}}^1(t)\psi(z) = \frac{\mathrm{d}\psi}{\mathrm{d}z}(z)\left[\Delta z + \frac{\mathrm{d}F}{\mathrm{d}u}(u_t)[z]\right] + \mathrm{Tr}\left[\frac{\mathrm{d}^2\psi}{\mathrm{d}z^2}(z)C(u_t)\right].$$

The process z is uniquely determined as the solution of the following martingale problem. For any measurable, bounded and twice Fréchet differentiable function $\psi : L^2(I) \to \mathbb{R}$, the process

$$\bar{N}_\psi(t) := \psi(z_t) - \int_0^t \bar{\mathcal{G}}^1(s)\psi(z_s)\mathrm{d}s \tag{4.29}$$

for $t \in [0,T]$, is a martingale.

The evolution equation associated to the martingale problem (4.29) is the following SPDE (see [DPZ92])

$$\mathrm{d}z_t = \left(\Delta z_t + \frac{\mathrm{d}F}{\mathrm{d}u}(u_t)[z_t]\right)\mathrm{d}t + \Gamma(u_t)\mathrm{d}W_t \tag{4.30}$$

with initial condition 0 and zero Dirichlet boundary conditions. The operator $\Gamma(u)$ is the square root of $C(u)$: $C(u) = \Gamma(u)\Gamma(u)^*$, which is well defined by Proposition 4.5.7. W denotes the standard cylindrical Wiener process on the Hilbert space $L^2(I)$. Formally, the cylindrical Wiener process W is defined as follows: let $((\beta_k(t))_{t\geq 0}, k \geq 1)$ be a family of independent Brownian motions, then

$$W_t = \sum_{k\geq 1} \beta_k(t) f_k,$$

where the definition of the Hilbert basis $\{f_k, k \geq 1\}$ of $L^2(I)$ is recalled in Chapter 2, Section 2.1. See [DPZ92], Sections 2.2.3 and 3.6 for more information about the construction of W. A complete description of the diffusion operator C is provided is Section 4.5.3. For any $u \in \mathcal{C}([0,T], L^2(I))$ and $t > 0$, the operator

$$Q_t : \psi \mapsto \int_0^t e^{\Delta(t-s)} C(u_s) e^{\Delta(t-s)} \psi \mathrm{d}s$$

is of trace class in $L^2(I)$ (Proposition 4.5.8). Thus we can apply classical results from the theory of SPDE in Hilbert spaces to deduce the existence and uniqueness of a mild solution to equation (4.30), see the classical reference [DPZ92] on this topic. The Langevin approximation of u

$$\mathrm{d}\tilde{u}^\varepsilon = [\Delta \tilde{u}^\varepsilon + F(\tilde{u}^\varepsilon)]\mathrm{d}t + \sqrt{\varepsilon}\Gamma(\tilde{u}^\varepsilon)\mathrm{d}W_t \tag{4.31}$$

is then well defined as stated in Proposition 4.5.9. It can be used as a tractable approximation of u^ε for small ε. For the order of convergence of $\tilde{u}^\varepsilon$ towards u, we refer to [EK86], Chapter 11, Section 3.

Theorem 4.4.1 extends to the multiscale case. In this case there are at least two classes E_j as described in Section 4.3.

Theorem 4.4.2. *When ε goes to 0, the process z^ε converges in distribution in $\mathcal{C}([0,T], L^2(I))$ towards a process z uniquely defined as the solution of the following martingale problem: for any measurable, bounded and twice Fréchet differentiable function $\psi : L^2(I) \to \mathbb{R}$, the process*

$$\bar{N}_\psi(t) := \psi(z_t) - \int_0^t \bar{\mathcal{G}}^1(u_s, \bar{r}_s)\psi(z_s)\mathrm{d}s \tag{4.32}$$

is a martingale for $t \in [0,T]$. The operator $\bar{\mathcal{G}}^1$ is given by

$$\bar{\mathcal{G}}^1(u,\bar{r})\psi(z) = \frac{\mathrm{d}\psi}{\mathrm{d}z}(z)\left[\Delta z + \frac{\mathrm{d}F_{\bar{r}}}{\mathrm{d}u}(u_t)[z]\right] + \mathrm{Tr}\left[\frac{\mathrm{d}^2\psi}{\mathrm{d}z^2}(z)C_{\bar{r}}(u)\right] \tag{4.33}$$

Note that the evolution of the limit process z is coupled with the evolution of $(\bar{r}_t, t \in [0,T])$ and $(u_t, t \in [0,T])$ in contrary to Theorem 4.4.1 where no jumps remain. The diffusion operator $C_{\bar{r}}(u) : L^2(I) \to L^2(I)$ is characterized by the quantities $(C_{\bar{r}}(u)f_j, f_i)_{L^2(I)}$ which are given by

$$\int_{\mathcal{R}} (G_r(u) - F_{\bar{r}}(u), f_i)_{L^2(I)}(\Phi(u,r), f_j)_{L^2(I)} \otimes_{i=1}^N \mu_{\bar{r}(i)}(u)(\mathrm{d}r),$$

Moreover $\Phi : L^2(I) \times \mathcal{R} \to L^2(I)$ is the unique solution of

$$\left\{ \begin{array}{rcll} \mathcal{B}(u)\Phi(u,r) & = & -(G_r(u) - F_{\bar{r}}(u)), & \forall (u,r) \in L^2(I) \times \mathcal{R} \\ \int_{\mathcal{R}} \Phi(u,r) \otimes_{i=1}^N \mu_{j_i}(u)(\mathrm{d}r) & = & 0, & \forall (j_1, \cdots, j_N) \in \{1, \ldots, l\}^N, \end{array} \right. \tag{4.34}$$

where $\mathcal{B}$ is the "fast" generator introduced in (4.19).

The evolution equation associated to the martingale problem (4.32) is no longer an SPDE but a hybrid SPDE satisfying

$$\mathrm{d}z_t = \left(\Delta z_t + \frac{\mathrm{d}F_{\bar{r}_t}}{\mathrm{d}u}(u_t)[z_t]\right)\mathrm{d}t + \Gamma_{\bar{r}_t}(u_t)\mathrm{d}W_t \tag{4.35}$$

with initial condition 0 and zero Dirichlet boundary conditions. For $(u, \bar{r})$ held fixed, $\Gamma_{\bar{r}}(u)$ is the square root of $C_{\bar{r}}(u)$: $C_{\bar{r}}(u) = \Gamma_{\bar{r}}(u)\Gamma_{\bar{r}}(u)^*$. Hence, two noise sources are present in the multiscale case: the ionic channel noise represented by the random jumps of the process $\bar{r}$ and the Gaussian noise due to the fluctuations induced by the white noise W. In between each jump of the component $\bar{r}$, the process z follows a classical SPDE parametrized by the current value of the process $\bar{r}$. The hybrid SPDE (4.35) is well defined if for each $\bar{r} = j \in \{1, \cdots, l\}$ held fixed, the SPDE

$$\mathrm{d}z_t = \left(\Delta z_t + \frac{\mathrm{d}F_j}{\mathrm{d}u}(u_t)[z_t]\right)\mathrm{d}t + \Gamma_j(u_t)\mathrm{d}W_t$$

is well defined. For any $(j, u) \in \{1, \cdots, l\} \times \mathcal{C}([0,T], L^2(I))$ and $t > 0$, one can show that the operator

$$Q_t^j : \psi \mapsto \int_0^t e^{\Delta(t-s)} C_j(u_s) e^{\Delta(t-s)} \psi \mathrm{d}s$$

is of trace class in $L^2(I)$. This allows us to apply classical results from the theory of SPDE in Hilbert spaces to deduce existence and uniqueness of a mild solution to equation (4.35). See also [YZ09] for an introduction to switching diffusions.

Theorem 4.4.1 (*all-fast* case) is proved in full details in Section 4.5. The proof of Theorem 4.4.2 (*multiscale* case) follows the same structure with an additional complication in the notations and the following necessary adaptations. Regarding the proof of tightness, the argument in Section 4.5.1 below relies on the Poisson equation. We refer the reader to Chapter 3 Section 3.3.2, which explains how the Poisson equation may be extended to the multiscale setting. Regarding the identification of the limit, we adapt the method of [Wai10], Chapter 5, Section 4.3, where the multiscale case is considered in the finite dimensional setting. The key point is to be able to write down the generator of the process $(z^\varepsilon, u^\varepsilon, \bar{r}^\varepsilon)$. For another instructive example dealing with slow-fast continuous Markov chain, see [YZ98], Chapter 7.

4.5 Proofs

In Theorem 4.4.1, we want to prove the convergence in distribution of the process z^ε when ε goes to zero. As usual in this context such a proof can be divided in two parts: the proof of tightness of the family $\{z^\varepsilon, \varepsilon \in]0,1]\}$ which implies that there exists a convergent subsequence and the identification of the limit which allows us to characterize the limit of any converging subsequence and prove its uniqueness. We write in full details the proof in the all fast case corresponding to Section 4.4.1, that is when all the states in E communicate at fast rates of order $\frac{1}{\varepsilon}$. In this case there is a unique class of fast communications which is the whole state space E (that is $l = 1$ w.r.t. the notation of section 4.3). As already noticed, the multiscale case (when $l > 1$) considered in Theorem 4.4.2 may be deduced from the all fast case and amounts mainly in additional complication in the notations.

4.5.1 Tightness

To show that the family $\{z^\varepsilon, \varepsilon \in]0,1]\}$ is tight in $\mathbb{D}([0,T], L^2(I))$, we use Aldous criterion (cf. [Mé84], or Chapter 2, Section 2.3) which can be splitted in two parts as follows.

Criterion 4.5.1 (General criterion for tightness [Mé84])**.** *Let us assume that the family* $\{z^\varepsilon, \varepsilon \in]0,1]\}$ *satisfies Aldous's condition: for any* $\delta, M > 0$*, there exist* $\eta, \varepsilon_0 > 0$ *such that for all stopping times* τ *with* $\tau + \eta < T$*,*

$$\sup_{\varepsilon \in]0,\varepsilon_0]} \sup_{\theta \in]0,\eta[} \mathbb{P}(\|z^\varepsilon_{\tau+\theta} - z^\varepsilon_\tau\|_{L^2(I)} \geq M) \leq \delta \tag{4.36}$$

and moreover, for each $t \in [0,T]$*, the family* $\{z^\varepsilon_t, \varepsilon \in]0,1]\}$ *is tight in* $L^2(I)$*. Then* $\{z^\varepsilon, \varepsilon \in]0,1]\}$ *is tight in* $\mathbb{D}([0,T], L^2(I))$*.*

Criterion 4.5.2 (Tightness in a Hilbert space [Mé84])**.** *Let* $L^2(I)$ *be a separable Hilbert space endowed with a basis* $\{f_k, k \geq 1\}$ *and for* $k \geq 1$ *define*

$$L_k = span\{f_i, 1 \leq i \leq k\}.$$

Then, for t held fixed, $(z^\varepsilon_t, \varepsilon \in]0,1])$ *is tight in* $L^2(I)$ *if, and only if, for any* $\delta, \eta > 0$ *there exist* $\rho, \varepsilon_0 > 0$ *and* $L_{\delta,\eta} \subset \{L_k, k \geq 1\}$ *such that*

$$\sup_{\varepsilon \in]0,\varepsilon_0]} \mathbb{P}(\|z^\varepsilon_t\|_{L^2(I)} > \rho) \leq \delta, \tag{4.37}$$

$$\sup_{\varepsilon \in]0,\varepsilon_0]} \mathbb{P}(d(z^\varepsilon_t, L_{\delta,\eta}) > \eta) \leq \delta, \tag{4.38}$$

where $d(z^\varepsilon_t, L_{\delta,\eta}) = \inf_{v \in L_{\delta,\eta}} \|z^\varepsilon_t - v\|_{L^2(I)}$ *is the distance of* z^ε *to the subspace* $L_{\delta,\eta}$*.*

We begin by showing that for a fixed $t \in [0,T]$, the family $\{z^\varepsilon_t, \varepsilon \in]0,1]\}$ is uniformly bounded in $L^2(\Omega, L^2(I))$. We recall the definition of the Hilbert basis $\{e_k, k \geq 1\}$ of $L^2(I)$ from Chapter 2, Section 2.1

$$f_k(x) = \sqrt{2}\sin(k\pi x), \quad x \in I.$$

Proposition 4.5.1. *There exists a constant C depending only on T but otherwise neither on $t \in [0,T]$ nor on $\varepsilon \in]0,1]$ such that*

$$\mathbb{E}(\|z_t^\varepsilon\|_{L^2(I)}^2) \leq C$$

In particular, for any fixed $t \in [0,T]$, condition (4.37) is satisfied by the family $\{z_t^\varepsilon, \varepsilon \in]0,1]\}$.

Proof. Let $t \in [0,T]$ and $\varepsilon \in]0,1]$ be fixed. Using the evolution equations on u^ε and u and plugging F given by (4.26) in the calculation, we have:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|u_t^\varepsilon - u_t\|_{L^2(I)}^2 &= 2 < \partial_t(u_t^\varepsilon - u_t), u_t^\varepsilon - u_t > \\
&= 2 < \Delta(u_t^\varepsilon - u_t), u_t^\varepsilon - u_t > +2 < G_{r_t^\varepsilon}(u_t^\varepsilon) - F(u_t), u_t^\varepsilon - u_t > \\
&= -2\|D(u_t^\varepsilon - u_t)\|_{L^2(I)}^2 + 2(G_{r_t^\varepsilon}(u_t^\varepsilon) - F(u_t^\varepsilon), u_t^\varepsilon - u_t)_{L^2(I)} \\
&\quad + 2(F(u_t^\varepsilon) - F(u_t), u_t^\varepsilon - u_t)_{L^2(I)},
\end{aligned}$$

almost surely. We treat each of the above terms separately. Regarding the third term, we notice that the application $u \mapsto (F(u), u)_{L^2(I)}$ is locally Lipschitz on $L^2(I)$ and that the quantities u_t^ε and u_t are uniformly bounded w.r.t. $t \in [0,T]$ and $\varepsilon \in]0,1]$ thanks to Propositions 4.3.1 and 4.3.2. Thus there exists a constant C, depending only on T but otherwise not on $t \in [0,T]$ and $\varepsilon \in]0,1]$, such that

$$2(F(u_t^\varepsilon) - F(u_t), u_t^\varepsilon - u_t)_{L^2(I)} \leq C\|u_t^\varepsilon - u_t\|_{L^2(I)}^2.$$

Integrating over $[0,t]$ and taking expectation yields the following inequality

$$\begin{aligned}
\mathbb{E}(\|u_t^\varepsilon - u_t\|_{L^2(I)}^2) &\leq \mathbb{E}(\|u_0^\varepsilon - u_0\|_{L^2(I)}^2) + 2C\int_0^t \mathbb{E}(\|u_s^\varepsilon - u_s\|_{L^2(I)}^2)\mathrm{d}s \\
&\quad + \mathbb{E}\left(\int_0^t 2(G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), u_s^\varepsilon - u_s)_{L^2(I)}\mathrm{d}s\right).
\end{aligned}$$

Let us consider the latter of these terms. Using the same approach as the one developed for the identification of the limit in the proof of the averaging result in [GT12], we deduce the existence of a constant $C(T)$ depending only on T such that

$$\left|\mathbb{E}\left(\int_0^t 2(G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), u_s^\varepsilon - u_s)_{L^2(I)}\mathrm{d}s\right)\right| \leq C(T)\varepsilon.$$

For the sake of completeness, we review now briefly this approach and refer to Chapter 3 for more details. Remark that, due to the regularity of the reaction term (4.6) which is in $L^2(I)$ and not only in $H^{-1}(I)$ as in Chapter 3, the situation is easier to handle. The key point is to show, as in Proposition 3.3.1 of Chapter 3, that there exists a measurable and bounded function $f : L^2(I) \times \mathcal{R} \times [0,T] \to \mathbb{R}$ such that $\int_{\mathcal{R}} f(u,r,t)\mu(u)(dr) = 0$ and for all $(u,r,t) \in L^2(I) \times \mathcal{R} \times [0,T]$

$$\mathcal{B}(u)f(u,\cdot,t)(r) = (G_r(u) - F(u), u - u_t)_{L^2(I)}. \tag{4.39}$$

Equation (4.39) is called the Poisson equation related to $\mathcal{B}$. Then using the regularity of the mappings $(u,r,t) \in L^2(I) \times \mathcal{R} \times [0,T] \mapsto (G_r(u) - F(u), u - u_t)_{L^2(I)}$ and the operator $\mathcal{B}(u)$ for $u \in L^2(I)$, we deduce that the application $(u,r,t) \in L^2(I) \times \mathcal{R} \times [0,T] \mapsto f(u,r,t)$ is bounded, Fréchet differentiable in u with bounded Fréchet derivative and differentiable in t with bounded derivative. Using the general theory of Markov processes, we deduce that there exists a martingale M^ε such that

$$\begin{aligned}
&f(u_t^\varepsilon, r_t^\varepsilon, t)\\
&= f(u_0^\varepsilon, r_0^\varepsilon, 0) \quad + \int_0^t \mathcal{A}^\varepsilon f(u_s^\varepsilon, r_s^\varepsilon, s)\mathrm{d}s + M_t^\varepsilon\\
&= f(u_0^\varepsilon, r_0^\varepsilon, 0) \quad + \frac{1}{\varepsilon}\int_0^t \mathcal{B}(u_s^\varepsilon)f(u_s^\varepsilon, r_s^\varepsilon, s) + \frac{\mathrm{d}f}{\mathrm{d}s}(u_\cdot^\varepsilon, r_s^\varepsilon, s)(s) + \frac{\mathrm{d}f}{\mathrm{d}s}(u_s^\varepsilon, r_s^\varepsilon, \cdot)(s)\mathrm{d}s + M_t^\varepsilon\\
&= f(u_0^\varepsilon, r_0^\varepsilon, 0) \quad + \frac{1}{\varepsilon}\int_0^t (G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), u_s^\varepsilon - u_s)_{L^2(I)}\mathrm{d}s\\
&\qquad\qquad\qquad + \int_0^t \frac{\mathrm{d}f}{\mathrm{d}s}(u_\cdot^\varepsilon, r_s^\varepsilon, s)(s) + \frac{\mathrm{d}f}{\mathrm{d}s}(u_s^\varepsilon, r_s^\varepsilon, \cdot)(s)\mathrm{d}s + M_t^\varepsilon.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\int_0^t (G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), u_s^\varepsilon - u_s)_{L^2(I)}\mathrm{d}s\\
&= \ \varepsilon f(u_t^\varepsilon, r_t^\varepsilon, t) - \varepsilon f(u_0^\varepsilon, r_0^\varepsilon, 0) - \varepsilon \int_0^t \frac{\mathrm{d}f}{\mathrm{d}s}(u_\cdot^\varepsilon, r_s^\varepsilon, s)(s) - \frac{\mathrm{d}f}{\mathrm{d}s}(u_s^\varepsilon, r_s^\varepsilon, \cdot)(s)\mathrm{d}s - \varepsilon M_t^\varepsilon.
\end{aligned}$$

Taking the expectation, using the fact that M^ε is a martingale and that f is regular, we obtain the desired estimate.
Assembling all the above estimates we obtain

$$\mathbb{E}(\|u_t^\varepsilon - u_t\|_{L^2(I)}^2) \ \leq \ \mathbb{E}(\|u_0^\varepsilon - u_0\|_{L^2(I)}^2) + C(T)\varepsilon + 2C\int_0^t \mathbb{E}(\|u_s^\varepsilon - u_s\|_{L^2(I)}^2)\mathrm{d}s.$$

Since $u_0^\varepsilon = u_0$ a standard application of Gronwall's lemma leads to the desired result. We end this proof by showing that for any fixed $t \in [0,T]$, the family $\{z_t^\varepsilon, \varepsilon \in]0,1]\}$ fulfills the requirement (4.37). Indeed, let $\delta > 0$ and denote by C the constant independent of ε and $t \in [0,T]$ such that

$$\mathbb{E}(\|z_t^\varepsilon\|_{L^2(I)}^2) \leq C.$$

By the Markov inequality we have, for $\rho > 0$,

$$\sup_{\varepsilon\in]0,1]} \mathbb{P}(\|z_t^\varepsilon\|_{L^2(I)} > \rho) \leq \sup_{\varepsilon\in]0,1]} \frac{\mathbb{E}(\|z_t^\varepsilon\|_{L^2(I)}^2)}{\rho^2} \leq \frac{C}{\rho^2}$$

and for ρ large enough, we obtain that $\sup_{\varepsilon\in]0,1]} \mathbb{P}(\|z_t^\varepsilon\|_{L^2(I)} > \rho) < \delta$. $\square$

We now prove the tightness of the family $\{z_t^\varepsilon, \varepsilon \in]0,1]\}$ in $L^2(I)$ for any fixed $t \in [0,T]$. This is the object of the following propositions.

Proposition 4.5.2. *Let $t \in]0,T]$ and for $p \geq 1$ let us define the following truncation*

$$z_t^{\varepsilon,p} = \sum_{k=1}^{p} (z_t^\varepsilon, f_k) f_k.$$

Then

$$\lim_{p\to\infty} \mathbb{E}(\|z_t^\varepsilon - z_t^{\varepsilon,p}\|_{L^2(I)}^2) = 0,$$

uniformly in $\varepsilon \in]0,1]$.

Proof. For a fixed $k \geq 1$ we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}(z_t^\varepsilon, e_k)^2 &= 2(z_t^\varepsilon, f_k)\frac{\mathrm{d}}{\mathrm{d}t}(z_t^\varepsilon, f_k) \\
&= 2(z_t^\varepsilon, f_k)\left(-(k\pi)^2(z_t^\varepsilon, f_k) + \frac{1}{\sqrt{\varepsilon}} < G_{r_t^\varepsilon}(u_t^\varepsilon) - F(u_t), f_k >\right) \\
&= -2(k\pi)^2(z_t^\varepsilon, f_k)^2 + \frac{2}{\sqrt{\varepsilon}}(z_t^\varepsilon, f_k)(F(u_t^\varepsilon) - F(u_t), f_k)_{L^2(I)} \\
&\quad + \frac{2}{\sqrt{\varepsilon}}(z_t^\varepsilon, f_k)(G_{r_t^\varepsilon}(u_t^\varepsilon) - F(u_t^\varepsilon), f_k)_{L^2(I)},
\end{aligned}$$

almost surely. A direct computation using the arguments developed in the proof of Proposition 4.5.1 leads to the existence of a constant $C(T)$ independent of $\varepsilon \in]0,1]$ such that

$$(z_t^\varepsilon, f_k)^2 \leq C(T) - 2(k\pi)^2 \int_0^t (z_s^\varepsilon, f_k)^2 \mathrm{d}s,$$

almost surely. Using Gronwall's lemma we deduce that

$$(z_t^\varepsilon, f_k)^2 \leq C(T)e^{-2(k\pi)^2 t}.$$

The result follows since the series $\sum e^{-2(k\pi)^2 t}$ is convergent for $t > 0$. $\square$

We now check that the family $\{z^\varepsilon, \varepsilon \in]0,1]\}$ satisfies the first part of Criterion 4.5.1.

Proposition 4.5.3. *Let $\tau > 0$ be a stopping time and $\theta > 0$ such that $\tau + \theta \le T$. There exists a constant C depending only on T such that*

$$\mathbb{E}(\|z^\varepsilon_{\tau+\theta} - z^\varepsilon_\tau\|^2_{L^2(I)}) \le C\theta.$$

Proof. We notice that for $k \ge 1$, $t > 0$ and $\theta > 0$ such that $t + \theta \le T$ we have

$$\partial_\theta (z^\varepsilon_{t+\theta} - z^\varepsilon_t, f_k)_{L^2(I)} = -(k\pi)^2 (z^\varepsilon_{t+\theta}, f_k)_{L^2(I)} + \frac{1}{\sqrt{\varepsilon}} (G_{r^\varepsilon_{t+\theta}}(u^\varepsilon_{t+\theta}) - F(u_{t+\theta}), f_k)_{L^2(I)}.$$

Thus, almost surely

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\theta}(z^\varepsilon_{t+\theta} - z^\varepsilon_t, f_k)^2_{L^2(I)} &= -2(k\pi)^2 (z^\varepsilon_{t+\theta}, f_k)_{L^2(I)} (z^\varepsilon_{t+\theta} - z^\varepsilon_t, f_k)_{L^2(I)} \\ &\quad + \frac{2}{\sqrt{\varepsilon}} < G_{r^\varepsilon_{t+\theta}}(u^\varepsilon_{t+\theta}) - F(u_{t+\theta}), f_k > (z^\varepsilon_{t+\theta} - z^\varepsilon_t, f_k)_{L^2(I)}. \end{aligned}$$

The first term satisfies

$$\begin{aligned} &- 2(k\pi)^2 (z^\varepsilon_{t+\theta}, f_k)_{L^2(I)} (z^\varepsilon_{t+\theta} - z^\varepsilon_t, f_k)_{L^2(I)} \\ &= -2(k\pi)^2 (z^\varepsilon_{t+\theta} - z^\varepsilon_t, f_k)^2 + 2(k\pi)^2 (z^\varepsilon_t, f_k)^2_{L^2(I)} - 2(k\pi)^2 (z^\varepsilon_{t+\theta}, f_k)_{L^2(I)} (z^\varepsilon_t, f_k)_{L^2(I)} \\ &\le -2(k\pi)^2 (z^\varepsilon_{t+\theta} - z^\varepsilon_t, f_k)^2 + 3(k\pi)^2 \|z^\varepsilon_t\|^2_{L^2(I)} + (k\pi)^2 \|z^\varepsilon_{t+\theta}\|^2_{L^2(I)} \end{aligned}$$

where $\|z^\varepsilon_t\|^2_{L^2(I)}$ and $\|z^\varepsilon_{t+\theta}\|^2_{L^2(I)}$ are bounded in expectation by a constant independent of t, θ and ε by Proposition 4.5.1. For the second term, the arguments developed in the proof of Proposition 4.5.1 lead to the existence of a constant C depending only on T such that

$$\mathbb{E}\left(\int_0^\theta (G_{r^\varepsilon_{t+s}}(u^\varepsilon_{t+s}) - F(u^\varepsilon_{t+s}), f_k)_{L^2(I)} \mathrm{d}s \right) \le C\theta\varepsilon.$$

Therefore, still denoting by C a constant depending only of T

$$\mathbb{E}((z^\varepsilon_{t+\theta} - z^\varepsilon_t, f_k)^2_{L^2(I)}) \le -2(k\pi)^2 \int_0^\theta \mathbb{E}((z^\varepsilon_{t+s} - z^\varepsilon_t, f_k)^2) \mathrm{d}s + C(1 + (k\pi)^2)\theta.$$

By application of the Gronwall's lemma and summation over k we obtain

$$\mathbb{E}(\|z^\varepsilon_{t+\theta} - z^\varepsilon_t\|^2_{L^2(I)}) \le C\theta \sum_{k\ge1} (1 + (k\pi)^2) e^{-2(k\pi)^2 t}$$

which yields the result for any $t > 0$ since the series $\sum_{k\ge1}(1 + (k\pi)^2)e^{-2(k\pi)^2 t}$ is convergent for $t > 0$. The same arguments apply when replacing t by the stopping time τ. □

According to Criteria 4.5.1 and 4.5.2, Propositions 4.5.1, 4.5.2 and 4.5.3, the family $\{z^\varepsilon, \varepsilon \in]0, 1]\}$ is tight in $\mathbb{D}([0, T], L^2(I))$. The continuity of each element of the family implies that $\{z^\varepsilon, \varepsilon \in]0, 1]\}$ is tight in $\mathcal{C}([0, T], L^2(I))$.

4.5.2 Identification of the limit

In this section we want to prove that $(z^\varepsilon, \varepsilon \in]0,1])$ has a unique accumulation point that we identify as the unique solution of a martingale problem. For this purpose, we study the process $(z^\varepsilon, r^\varepsilon)$ for $\varepsilon \in]0,1]$.
Let us outline the strategy of the proof.

Step 1. Use the general theory on PDMP developed in [BR11] to write down the generator $\mathcal{G}^\varepsilon$ of the process $(z^\varepsilon, r^\varepsilon)$. The associated martingale problem gives rise to martingales M^ε_ϕ for appropriate functions ϕ.

Step 2. For a nice choice of ϕ, identify the terms of order one in ε of the martingale M^ε_ϕ. Since the difference between u and u^ε is renormalized by $\sqrt{\varepsilon}$, choose ϕ of the form $\psi + \sqrt{\varepsilon}\gamma$ (perturbed test function).

Step 3. Identify the generator $\bar{\mathcal{G}}$ of the limit process z. Prove that z is solution of the martingale problem associated to $\bar{\mathcal{G}}$.

Step 1. Notice first that the process z^ε satisfies the following equation

$$\begin{aligned} \partial_t z^\varepsilon_t &= \Delta z^\varepsilon_t + \frac{1}{\sqrt{\varepsilon}}(G_{r^\varepsilon_t}(u^\varepsilon_t) - F(u_t)) \qquad (4.40) \\ &= \Delta z^\varepsilon_t + \frac{1}{\sqrt{\varepsilon}}(G_{r^\varepsilon_t}(u_t + \sqrt{\varepsilon} z^\varepsilon_t) - F(u_t)), \end{aligned}$$

by definition of z^ε. The initial condition for z^ε is 0 and the boundary conditions are still zero Dirichlet boundary conditions.
Let $\phi : L^2(I) \times \mathcal{R} \times \mathbb{R}_+$ be a real valued, measurable and bounded function of class $\mathcal{C}^2$ on $L^2(I)$ and $\mathcal{C}^1$ on $\mathbb{R}_+$. We write down the generator of the process $(z^\varepsilon, r^\varepsilon)$ against ϕ. Recall that in the all-fast case, the limit u of u^ε is deterministic so that $(z^\varepsilon, r^\varepsilon)$ is a classical PDMP with evolution equation given by (4.40) and dynamic of jumps given by (4.7). According to Theorem 4 of [BR11], for $(z, r, t) \in L^2(I) \times \mathcal{R} \times \mathbb{R}_+$, the generator $\mathcal{G}$ of $(z^\varepsilon, r^\varepsilon)$ is given by

$$\begin{aligned} \mathcal{G}(t)\phi(z,r,t) =& \frac{\mathrm{d}\phi}{\mathrm{d}z}(z,r,t)[\Delta z + \frac{1}{\sqrt{\varepsilon}}(G_r(u_t + \sqrt{\varepsilon} z) - F(u_t))] \qquad (4.41) \\ &+ \frac{1}{\varepsilon}\mathcal{B}(u_t + \sqrt{\varepsilon} z)\phi(z,r,t) + \partial_t \phi(z,r,t). \end{aligned}$$

Following the usual theory of Markov processes, see [EK86], Chapter 4, the process $(M^\varepsilon_\phi(t), t \in [0,T])$ defined for $t \geq 0$ by

$$M^\varepsilon_\phi(t) = \phi(z^\varepsilon_t, r^\varepsilon_t, t) - \int_0^t \mathcal{G}(s)\phi(z^\varepsilon_s, r^\varepsilon_s, s)\mathrm{d}s,$$

is a martingale for the natural filtration associated to the process $(z^\varepsilon, r^\varepsilon)$.

Step 2. We want to identify the terms of different orders in ε of M^ε_ϕ. For this purpose, we choose a function ϕ with the following decomposition

$$\phi(z,r,t) = \psi(z,r) + \sqrt{\varepsilon}\gamma(z,r,t),$$

where the functions ψ and γ have the same regularity as ϕ. We write the Taylor expansion in ε of the two following terms

$$\begin{aligned} G_r(u_t + \sqrt{\varepsilon}z) &= G_r(u_t) + \sqrt{\varepsilon}\frac{\mathrm{d}G_r}{\mathrm{d}u}(u_t)[z] + \sqrt{\varepsilon}\|z\|_{L^2(I)}\delta_1(\sqrt{\varepsilon}z) \\ \mathcal{B}(u_t + \sqrt{\varepsilon}z) &= \mathcal{B}(u_t) + \sqrt{\varepsilon}\frac{\mathrm{d}\mathcal{B}}{\mathrm{d}u}(u_t)[z] + \sqrt{\varepsilon}\|z\|_{L^2(I)}\delta_2(\sqrt{\varepsilon}z), \end{aligned}$$

where δ_1 and δ_2 are two $L^2(I)$-valued continuous functions such that $\delta_1(0_{L^2(I)}) = \delta_2(0_{L^2(I)}) = 0_{L^2(I)}$. Plugging this expansion in the expression of the generator (4.41) we want the terms of order $\frac{1}{\varepsilon}$ to vanish. For $(z,r,t) \in L^2(I) \times \mathcal{R} \times \mathbb{R}_+$ this leads to

$$\mathcal{B}(u_t)\psi(z,r) = 0. \tag{4.42}$$

That is to say, the application ψ does not depend on $r \in \mathcal{R}$ and is of the form

$$\psi(z,r) = \psi(z),$$

where $\psi : L^2(I) \to \mathbb{R}$ is of class $\mathcal{C}^2$. The generator is then of the following form, where we gather the terms of the same order in ε

$$\begin{aligned} &\mathcal{G}(t)\phi(z,r,t) \\ &= \frac{1}{\sqrt{\varepsilon}}\left(\frac{\mathrm{d}\psi}{\mathrm{d}z}(z)[G_r(u_t) - F(u_t)] + \mathcal{B}(u_t)\gamma(z,r,t) + \frac{\mathrm{d}\mathcal{B}}{\mathrm{d}u}(u_t)[z]\psi(z)\right) \\ &+ \frac{\mathrm{d}\psi}{\mathrm{d}z}(z)\left[\Delta z + \frac{\mathrm{d}G_r}{\mathrm{d}u}(u_t)[z]\right] + \frac{\mathrm{d}\gamma}{\mathrm{d}z}(z,r,t)[G_r(u_t) - F(u_t)] + \frac{\mathrm{d}\mathcal{B}}{\mathrm{d}u}(u_t)[z]\gamma(z,r,t) \\ &+ \sqrt{\varepsilon}\left(\partial_t\gamma(z,r,t) + \frac{\mathrm{d}\gamma}{\mathrm{d}z}(z,r,t)\left[\Delta z + \frac{\mathrm{d}G_r}{\mathrm{d}u}(u_t)[z]\right]\right) + \mathrm{o}(\sqrt{\varepsilon}). \end{aligned}$$

We now want the terms of order $\frac{1}{\sqrt{\varepsilon}}$ to vanish, that is to say, for $(z,r,t) \in L^2(I) \times \mathcal{R} \times \mathbb{R}_+$

$$\frac{\mathrm{d}\psi}{\mathrm{d}z}(z)[G_r(u_t) - F(u_t)] + \mathcal{B}(u_t)\gamma(z,r,t) + \frac{\mathrm{d}\mathcal{B}}{\mathrm{d}u}(u_t)[z]\psi(z) = 0.$$

Notice that $\mathcal{B}(u_t)1 = 0$ implies that for all $(z,t) \in L^2(I) \times \mathbb{R}_+$

$$\frac{\mathrm{d}\mathcal{B}}{\mathrm{d}u}(u_t)[z]\psi(z) = 0$$

and we are left with the equation

$$\mathcal{B}(u_t)\gamma(z,r,t) = -\frac{\mathrm{d}\psi}{\mathrm{d}z}(z)[G_r(u_t) - F(u_t)]. \tag{4.43}$$

We look for γ of the form:

$$\gamma(z,r,t) = \frac{\mathrm{d}\psi}{\mathrm{d}z}(z)[\Phi(r,u_t)],$$

where $\Phi : \mathcal{R} \times L^2(I) \to L^2(I)$ has to be identified. Inserting the above expression of γ in (4.43) we obtain

$$\frac{\mathrm{d}\psi}{\mathrm{d}z}(z)[\mathcal{B}(u_t)\Phi(r,u_t)] = -\frac{\mathrm{d}\psi}{\mathrm{d}z}(z)[G_r(u_t) - F(u_t)].$$

Therefore, it is enough that for any $(u,r) \in L^2(I) \times \mathcal{R}$

$$\mathcal{B}(u)\Phi(r,u) = -(G_r(u) - F(u)). \tag{4.44}$$

To ensure uniqueness of the solution for equation (4.44) we impose moreover the condition

$$\int_{\mathcal{R}} \Phi(r,u)\mu(u)(\mathrm{d}r) = 0.$$

Then, from the definition of F we have $\int_{\mathcal{R}} (G_r(u) - F(u))\, \mu(u)(dr) = 0$. Moreover, equation (4.44) has a unique solution Φ thanks to the Fredholm alternative.

Step 3. We have identified the terms of order 1 in ε of the generator of the process $(z^\varepsilon, r^\varepsilon)$. It remains to show that the terms of order 1 in ε correspond, after averaging, to the generator of the process z. For $(z,r,t) \in L^2(I) \times \mathcal{R} \times \mathbb{R}_+$ we define

$$\begin{aligned} \mathcal{G}^1(t,r)\psi(z) \;=\; & \frac{\mathrm{d}\psi}{\mathrm{d}z}(z)\left[\Delta z + \frac{\mathrm{d}G_r}{\mathrm{d}u}(u_t)[z]\right] + \frac{\mathrm{d}^2\psi}{\mathrm{d}z^2}(z)[\Phi(r,u_t), G_r(u_t) - F(u_t)] \\ & + \frac{\mathrm{d}\mathcal{B}}{\mathrm{d}u}(u_t)[z]\frac{\mathrm{d}\psi}{\mathrm{d}z}(z)[\Phi(r,u_t)]. \end{aligned} \tag{4.45}$$

Let us define also the following process

$$N_\psi^\varepsilon(t) = \psi(z_t^\varepsilon) - \int_0^t \mathcal{G}^1(s, r_s^\varepsilon)\psi(z_s^\varepsilon)\mathrm{d}s.$$

By construction we see that $\mathbb{E}(|M_\phi^\varepsilon(t) - N_\psi^\varepsilon(t)|^2) = \mathrm{O}(\varepsilon)$. When ε goes to 0, by the averaging result of Theorem 4.3.1, we see that the term $\int_0^t \mathcal{G}^1(s, r_s^\varepsilon)\psi(z_s^\varepsilon)ds$ should converge to

$$\int_0^t \int_{\mathcal{R}} \mathcal{G}^1(s,r)\psi(z_s)\mu(u_s)(dr)\mathrm{d}s.$$

Therefore, we want to prove that, whenever z is an accumulation point of the family $(z^\varepsilon, \varepsilon \in]0,1])$, the process

$$\bar{N}_\psi(t) = \psi(z_t) - \int_0^t \bar{\mathcal{G}}^1(s)\psi(z_s)\mathrm{d}s,$$

is a martingale w.r.t. the natural filtration associated to the process $(z_t, t \geq 0)$ where

$$\bar{\mathcal{G}}^1(t)\psi(z) = \frac{\mathrm{d}\psi}{\mathrm{d}z}(z)[\Delta z + \frac{\mathrm{d}F}{\mathrm{d}u}(u_t)[z]] + \frac{\mathrm{d}^2\psi}{\mathrm{d}z^2}(z)\int_{\mathcal{R}}[\Phi(r,u_t), G_r(u_t) - F(u_t)]\mu(u_t)(\mathrm{d}r). \tag{4.46}$$

This is not straightforward since we have no information on the asymptotic behavior of the process $(z^\varepsilon, r^\varepsilon)$ when ε goes to 0.

Proposition 4.5.4. *The process $(\bar{N}_\psi(t), t \geq 0)$ is a martingale w.r.t. the natural filtration associated to the process $(z_t, t \geq 0)$.*

Proof. Let $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k \leq s \leq t$ be $k+2$ reals, with $k \geq 1$ an integer. For $i \in \{1, \cdots, k\}$, we take a measurable and bounded function g_i. In order to show that the process $(\bar{N}_\psi(t), t \geq 0)$ is a martingale for the natural filtration associated to the process $(z_t, t \geq 0)$ we will prove that

$$\mathbb{E}((\bar{N}_\psi(t) - \bar{N}_\psi(s))g_1(z_{t_1}) \cdots g_k(z_{t_k})) = 0.$$

In order to not overload the proof with too many computations, we write Z_k for the random variable $g_1(z_{t_1}) \cdots g_k(z_{t_k})$ and Z_k^ε for $g_1(z_{t_1}^\varepsilon) \cdots g_k(z_{t_k}^\varepsilon)$. Using elementary substitution and the fact that z^ε converges in law toward z when ε goes to 0 we have

$$\begin{aligned}
&\mathbb{E}((\bar{N}_\psi(t) - \bar{N}_\psi(s))Z_k) \\
&= \mathbb{E}\left(\left(\psi(z_t) - \psi(z_s) - \int_s^t \bar{\mathcal{G}}^1(l)\psi(z_l)\mathrm{d}l\right) Z_k\right) \\
&= \mathbb{E}((\psi(z_t) - \psi(z_s))Z_k) - \mathbb{E}\left(\left(\int_s^t \bar{\mathcal{G}}^1(l)\psi(z_l)\mathrm{d}l\right) Z_k\right) \\
&= \lim_{\varepsilon\to 0} \mathbb{E}((\psi(z_t^\varepsilon) - \psi(z_s^\varepsilon))Z_k^\varepsilon) - \mathbb{E}\left(\left(\int_s^t \bar{\mathcal{G}}^1(l)\psi(z_l)\mathrm{d}l\right) Z_k\right) \\
&= \lim_{\varepsilon\to 0} \mathbb{E}((N_\psi^\varepsilon(t) - N_\psi^\varepsilon(s))Z_k^\varepsilon) + \lim_{\varepsilon\to 0} \mathbb{E}\left(\left(\int_s^t \mathcal{G}^1(l, r_l^\varepsilon)\psi(z_l^\varepsilon)dl\right) Z_k^\varepsilon\right) \\
&\quad - \mathbb{E}\left(\left(\int_s^t \bar{\mathcal{G}}^1(l)\psi(z_l)\mathrm{d}l\right) Z_k\right).
\end{aligned}$$

On one hand, from the definition of N_ψ^ε and the previous study of the different orders in ε of the martingale M_ϕ^ε we see that

$$\lim_{\varepsilon\to 0}\mathbb{E}((N_\psi^\varepsilon(t)-N_\psi^\varepsilon(s))Z_k^\varepsilon)=\lim_{\varepsilon\to 0}\mathbb{E}((N_\psi^\varepsilon(t)-N_\psi^\varepsilon(s))Z_k^\varepsilon)-\mathbb{E}((M_\phi^\varepsilon(t)-M_\phi^\varepsilon(s))Z_k^\varepsilon)$$

From the previous study of the different orders in ε, the right hand side is $\mathrm{O}(\sqrt{\varepsilon})$ and therefore converges to 0 when ε goes to 0. On the other hand, for $\varepsilon^1>0$ which will be chosen later

$$\lim_{\varepsilon\to 0}\mathbb{E}\left(\left(\int_s^t \mathcal{G}^1(l,r_l^\varepsilon)\psi(z_l^\varepsilon)dl\right)Z_k^\varepsilon\right)-\mathbb{E}\left(\left(\int_s^t \bar{\mathcal{G}}^1(l)\psi(z_l)\mathrm{d}l\right)Z_k\right) \tag{4.47}$$

$$=\lim_{\varepsilon\to 0}\mathbb{E}\left(\left(\int_s^t \mathcal{G}^1(l,r_l^\varepsilon)\psi(z_l^\varepsilon)dl\right)Z_k^\varepsilon\right)-\mathbb{E}\left(\left(\int_s^t \mathcal{G}^1(l,r_l^{\varepsilon^1})\psi(z_l^\varepsilon)\mathrm{d}l\right)Z_k^\varepsilon\right) \tag{4.48}$$

$$+\lim_{\varepsilon\to 0}\mathbb{E}\left(\left(\int_s^t \mathcal{G}^1(l,r_l^{\varepsilon^1})\psi(z_l^\varepsilon)dl\right)Z_k^\varepsilon\right)-\mathbb{E}\left(\left(\int_s^t \mathcal{G}^1(l,r_l^{\varepsilon^1})\psi(z_l)\mathrm{d}l\right)Z_k\right) \tag{4.49}$$

$$+\mathbb{E}\left(\left(\int_s^t \mathcal{G}^1(l,r_l^{\varepsilon^1})\psi(z_l)dl\right)Z_k\right)-\mathbb{E}\left(\left(\int_s^t \bar{\mathcal{G}}^1(l)\psi(z_l)\mathrm{d}l\right)Z_k\right). \tag{4.50}$$

We know that the quantity corresponding to (4.50) can be made arbitrarily small by conditioning appropriately (as in the proof of Proposition 4.5.1 for example) for small enough ε^1. Then, since z^ε converges in law towards z when ε goes to 0, the quantity (4.49) converges to 0 when ε goes to 0. This shows finally that (4.47) converges to 0 when ε goes to 0 and therefore

$$\mathbb{E}((\bar{N}_\psi(t)-\bar{N}_\psi(s))g_1(z_{t_1})\cdots g_k(z_{t_k}))=0,$$

as announced. □

We can now conclude that the limit process z is solution of the following martingale problem: for any measurable, bounded and twice Fréchet differentiable function ψ, the process defined by

$$\bar{N}_\psi(t)=\psi(z_t)-\int_0^t \bar{\mathcal{G}}^1(s)\psi(z_s)\mathrm{d}s,$$

for $t\in[0,T]$ is a martingale, where $\bar{\mathcal{G}}^1$ is given by (4.46).
In other word, the limit process z is solution to the martingale problem associated with the operator $\bar{\mathcal{G}}^1$. Then z is a solution of the SPDE (4.30) where the diffusion operator $C(u)$ for $u\in L^2(I)$ is identified thanks to the relation

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}z^2}(z)\int_{\mathcal{R}}[\Phi(r,u),G_r(u)-F(u)]\mu(u)(\mathrm{d}r)=\mathrm{Tr}\ \frac{\mathrm{d}^2\psi}{\mathrm{d}z^2}(z)C(u) \tag{4.51}$$

for $(u,z)\in L^2(I)\times L^2(I)$. The uniqueness of z follows from the properties of the Laplacian operator, the reaction term $\frac{\mathrm{d}F}{\mathrm{d}u}$ and the operator $C(u)$. For more insight in the properties of the diffusion operator, see the following section.

4.5.3 The diffusion operator C

In this section, we give more details about the diffusion operator C. In particular, we make explicit the dependence of Φ in (4.51) w.r.t. the data of our problem.

Proposition 4.5.5 (First representation of the diffusion operator)**.** *For $u \in L^2(I)$ and $r \in \mathcal{R}$ we have:*

$$\Phi(r,u) = -(\mu^*(u)\mu(u) + \mathcal{B}^*(u)\mathcal{B}(u))^{-1}\mathcal{B}^*(u)(G_\cdot(u) - F(u))(r).$$

That is, the function $\Phi(\cdot,u)$ is explicitly given as a function of the "fast jumping part" operator $\mathcal{B}(u)$ and the associated invariant measure $\mu(u)$.

Proof. The application Φ is defined by the two conditions

$$\left\{ \begin{array}{rcl} \mathcal{B}(u)\Phi(r,u) & = & -(G_r(u) - F(u)) \\ \int_{\mathcal{R}} \Phi(r,u)\mu(u)(\mathrm{d}r) & = & 0 \end{array} \right. \tag{4.52}$$

for $(u,r) \in L^2(I) \times \mathcal{R}$. Let $u \in L^2(I)$ be held fixed. Defining $D(u) = (\mu(u), \mathcal{B}(u))^T$ reduces (4.52) to

$$D(u)\Phi(\cdot,u) = -\begin{pmatrix} 0 \\ G_\cdot(u) - F(u) \end{pmatrix}.$$

Then

$$D^*(u)D(u)\Phi(\cdot,u) = -D^*(u)\begin{pmatrix} 0 \\ G_\cdot(u) - F(u) \end{pmatrix}.$$

It remains to prove that the operator $D^*(u)D(u)$ is invertible which is the key point to conclude. Indeed

$$D^*(u)D(u) = \mu(u)^*\mu(u) + \mathcal{B}^*(u)\mathcal{B}(u)$$

and the kernel of the two operators $\mu(u)^*\mu(u)$ and $\mathcal{B}^*(u)\mathcal{B}(u)$ are in direct sum and span the whole space $\mathbb{R}^{|\mathcal{R}|}$. Let $\mathrm{x} \in \mathrm{Ker}D^*(u)D(u)$, then x can be written uniquely as $\mathrm{z} + \mathrm{y}$ with $\mathrm{z} \in \mathrm{Ker}\mu(u)^*\mu(u)$ and $\mathrm{y} \in \mathrm{Ker}\mathcal{B}^*(u)\mathcal{B}(u)$. We have

$$\mu(u)^*\mu(u)\mathrm{y} + \mathcal{B}^*(u)\mathcal{B}(u)\mathrm{z} = 0.$$

Since $\mathcal{B}(u)\mathbf{1} = 0$ (where $\mathbf{1} \in \mathbb{R}^{|\mathcal{R}|}$) and $\mu(u)\mathbf{1} = 1$, multiplying the above equation to the left by $\mathbf{1}^T$ we have

$$\mu(u)\mathrm{y} = 0.$$

Since $y \in \mathrm{Ker}\mathcal{B}^*(u)\mathcal{B}(u) = \mathrm{Ker}\mathcal{B}(u) = \mathrm{span}\mathbf{1}$ here, we have $\mathrm{y} = y\mathbf{1}$ with $y \in \mathbb{R}$ and

$$\mu(u)\mathrm{y} = \mu(u)y\mathbf{1} = y$$

and thus $y = 0$ and $\mathrm{y} = 0$. Therefore $\mathrm{x} = \mathrm{z} \in \mathrm{Ker}\mu(u)^*\mu(u)$ and $\mathcal{B}^*(u)\mathcal{B}(u)\mathrm{z} = 0$. Thus $z \in \mathrm{Ker}\mu(u)^*\mu(u) \cap \mathrm{Ker}\mathcal{B}^*(u)\mathcal{B}(u) = \{0\}$ and $\mathrm{x} = \mathrm{z} = 0$. The operator $D^*(u)D(u)$ is then invertible. $\square$

Proposition 4.5.6 (Second representation of the diffusion operator)**.** *For any $(u,r) \in L^2(I) \times \mathcal{R}$*

$$\Phi(u,r) = \int_0^\infty \mathbb{E}_r(G_{r_s^u}(u) - F(u))\mathrm{d}s,$$

where for a given u, r^u denotes a Markov chain on $\mathcal{R}$ with transition rates $q_{r\tilde{r}}$ (c.f. (4.2)).

Proof. The process

$$M_t = \Phi(u, r_t^u) - \Phi(u,r) - \int_0^t \mathcal{B}(u)\Phi(u, r_s^u)\mathrm{d}s,$$

is a martingale w.r.t. the natural filtration generated by the process r^u. Let us take expectation and remember that

$$\left\{ \begin{array}{rcl} \mathcal{B}(u)\Phi(r,u) & = & -(G_r(u) - F(u)) \\ \int_{\mathcal{R}} \Phi(r,u)\mu(u)(\mathrm{d}r) & = & 0, \end{array} \right. . \tag{4.53}$$

Then,

$$\mathbb{E}_r(\Phi(u, r_t^u)) = \Phi(u,r) - \int_0^t \mathbb{E}_r(G_{r_s^u}(u) - F(u))\mathrm{d}s.$$

The desired result follows since:

$$\lim_{t\to\infty} \mathbb{E}_r(\Phi(u, r_t^u)) = \int_{\mathcal{R}} \Phi(r,u)\mu(u)(\mathrm{d}r) = 0.$$

□

Proposition 4.5.7. *The diffusion operator $C(u)$, for $u \in L^2(I)$, is positive in the sense that*

$$\mathrm{Tr}\ C(u) \geq 0.$$

Therefore the operator $\Gamma(u)$ such that $C(u) = \Gamma^*(u)\Gamma(u)$ is well defined.

Proof. For $u \in L^2(I)$ we have:

$$\begin{aligned} \mathrm{Tr}\ C(u) &= \sum_{k\geq 1} \int_{\mathcal{R}} (G_r(u) - F(u), f_k)_{L^2(I)}(\Phi(r,u), f_k)_{L^2(I)}\mu(u)(\mathrm{d}r) \\ &= -\sum_{k\geq 1} \int_{\mathcal{R}} (\mathcal{B}(u)\Phi(r,u), f_k)_{L^2(I)}(\Phi(r,u), f_k)_{L^2(I)}\mu(u)(\mathrm{d}r). \end{aligned}$$

We conclude that Tr $C(u) \geq 0$ because all the eigenvalues of the operator $\mathcal{B}(u)$ are non positive. □

Proposition 4.5.8. *The following estimate holds,*

$$\mathrm{Tr}\ \int_0^t e^{\Delta(t-s)}C(u_s)e^{\Delta(t-s)}\mathrm{d}s \leq \sum_{k\geq 1}\int_0^t (\alpha\|u_s\|^2_{L^2(I)} + \beta\|u_s\|_{L^2(I)} + \gamma)e^{-2(k\pi)^2(t-s)}\mathrm{d}s$$

for all $t \in [0,T]$ and all functions $u \in \mathcal{C}([0,T], L^2(I))$. The trace is taken in the $L^2(I)$-sense and α, β, γ are three constants.

Proof. This is a direct consequence of Proposition 4.5.6. Proposition 4.5.6 implies that

$$|(\Phi(u_s, r), f_k)_{L^2(I)}| \leq \frac{c_1}{N}\sum_{i=1}^N |(\phi_{z_i}, f_k)_{L^2(I)}|(1 + \|u_s\|_{L^2(I)})$$

for a constant c_1 and $(f_k, k \geq 1)$ a Hilbert basis of $L^2(I)$. Since each ϕ_{z_i} is in $L^2(I)$ we obtain

$$|(\Phi(u_s, r), f_k)_{L^2(I)}| \leq c_1(1 + \|u_s\|_{L^2(I)})$$

for another constant c_1. Let us write, in the same way as in the proof of Proposition 4.5.7

$$\begin{aligned} &\mathrm{Tr}\ \int_0^t e^{\Delta(t-s)}C(u_s)e^{\Delta(t-s)}\mathrm{d}s \\ &\leq \sum_{k\geq 1}\int_0^t e^{-2(k\pi)^2(t-s)}\int_{\mathcal{R}}(G_r(u_s) - F(u_s), f_k)_{L^2(I)}(\Phi(r, u_s), f_k)_{L^2(I)}\mu(u_s)(\mathrm{d}r)\mathrm{d}s. \end{aligned} \tag{4.54}$$

Using the explicit expression of $G_r(u) - F(u)$, it is not difficult to show that there exists a constant c_2 such that

$$|(G_r(u_s) - F(u_s), f_k)_{L^2(I)}| \leq c_2(1 + \|u_s\|_{L^2(I)}).$$

Plugging the latter inequality in (4.54) leads to the result. An explicit computation of Tr C is presented in Section 4.6. □

In particular, the operator Q_t defined by

$$Q_t : \psi \in\mapsto \int_0^t e^{\Delta(t-s)}C(u_s)e^{\Delta(t-s)}\psi\mathrm{d}s$$

with $(j, u) \in \{1, \cdots, l\} \times \mathcal{C}([0,T], L^2(I))$ is of trace class in $L^2(I)$. The Langevin approximation of u is then well defined as stated in the following proposition. We recall that in the all-fast case

$$F(u) = \frac{1}{N}\sum_{i=1}^N \sum_{\xi\in E} c_\xi \mu(\bar{u}_i)(\xi)(v_\xi - \bar{u}_i)\phi_{z_i}$$

for $u \in L^2(I)$.

Proposition 4.5.9. *Let $\varepsilon > 0$. The SPDE*

$$\mathrm{d}\tilde{u}^\varepsilon = [\Delta \tilde{u}^\varepsilon + F(\tilde{u}^\varepsilon)]\mathrm{d}t + \sqrt{\varepsilon}\Gamma(\tilde{u}^\varepsilon)\mathrm{d}W_t \tag{4.55}$$

with initial condition u_0 and zero Dirichlet boundary condition has a unique solution with sample paths in $\mathcal{C}([0,T], L^2(I))$. Moreover the quantity

$$\sup_{t\in[0,T]} \mathbb{E}(\|\tilde{u}_t^\varepsilon\|^2_{L^2(I)}) < \infty. \tag{4.56}$$

Proof. Thanks to the properties of the laplacian operator, the local Lipschitz continuity of F and Proposition 4.5.8, we can apply classical results on SPDE, see for example [DPZ92], Chapter 7, Theorem 7.4 to prove existence and uniqueness of solution to (4.55) in $\mathcal{C}([0,T], L^2(I))$. □

4.6 Example

We consider in this section a spatially extended stochastic Morris-Lecar model. Since the seminal work [ML81], the deterministic Morris-Lecar model is considered as one of the classical mathematical models for investigating neuronal behavior. At first, this model was designed to describe the voltage dynamic of the barnacle giant muscle fiber (see [ML81] for a complete description of the deterministic Morris-Lecar model). To take into account the intrinsic variability of the ion channels dynamic, a stochastic interpretation of this class of models has been introduced (see [BR11] and [Wai10], Chapter 5, Section 3) in which ion channels are modeled by jump Markov processes. The model then falls into the class of stochastic generalized Hodgkin-Huxley models considered in the present paper. Let us proceed to the mathematical description of the spatially extended stochastic Morris-Lecar model. In this model, the total current $G_{r^{\mathrm{K}},r^{\mathrm{Ca}}}(u)$ is given by

$$\frac{1}{C}\left[\frac{1}{N_{\mathrm{K}}}\sum_{i=1}^{N_{\mathrm{K}}} 1_1(r^{\mathrm{K}}(i))c_{\mathrm{K}}(v_{\mathrm{K}} - \bar{u}_i)\phi_{z_i} + \frac{1}{N_{\mathrm{Ca}}}\sum_{i=1}^{N_{\mathrm{Ca}}} 1_1(r^{\mathrm{Ca}}(i))c_{\mathrm{Ca}}(v_{\mathrm{Ca}} - \bar{u}_i)\phi_{z_i} + I\right]$$

and the evolution equation for the transmembrane potential

$$\partial_t u_t = \frac{a}{2RC}\Delta u_t + G_{r_t^{\mathrm{K}},r_t^{\mathrm{Ca}}}(u_t),$$

on $[0,T]\times[0,1]$ and with zero Dirichlet boundary condition. The total current is the sum of the potassium K current, the calcium Ca current and the impulse I. The positive constants a, R, C are relative to the bio-physical properties of the membrane. When the voltage is held fixed, for any $1 \le i \le N_{\mathrm{q}}$ where q is equal to K or Ca, $r^q(i)$ is a continuous time Markov chain with only two states 0 for *closed* and 1 for *open*. The jump rate from 1 to 0 is given by $\beta_q(\bar{u}_i)$ and from 0 to 1 by $\alpha_q(\bar{u}_i)$. All the jump rates are bounded below and above by positive constants. We will assume that the potassium ion channels communicate at fast rates of order $\frac{1}{\varepsilon}$ for a small $\varepsilon > 0$. The calcium rates are of order 1. The invariant measure associated to each channel $1 \leq i \leq N_{\mathrm{K}}$ is given by

$$\mu_i^{\mathrm{K}}(\bar{u}_i) = \left(\frac{\beta_{\mathrm{K}}(\bar{u}_i)}{\alpha_{\mathrm{K}}(\bar{u}_i) + \beta_{\mathrm{K}}(\bar{u}_i)}, \frac{\alpha_{\mathrm{K}}(\bar{u}_i)}{\alpha_{\mathrm{K}}(u(z_i)) + \beta_{\mathrm{K}}(\bar{u}_i)} \right).$$

Therefore the averaged applied current is

$$\begin{aligned} F_{r^{\mathrm{Ca}}}(u) &= \frac{1}{C}\left[\frac{1}{N_{\mathrm{K}}} \sum_{i=1}^{N_{\mathrm{K}}} \frac{\alpha_{\mathrm{K}}(\bar{u}_i)}{\alpha_{\mathrm{K}}(\bar{u}_i) + \beta_{\mathrm{K}}(\bar{u}_i)} c_{\mathrm{K}}(v_{\mathrm{K}} - \bar{u}_i)\phi_{z_i} \right. \\ &+ \left. \frac{1}{N_{\mathrm{Ca}}} \sum_{i=1}^{N_{\mathrm{Ca}}} 1_1(r^{\mathrm{Ca}}(i)) c_{\mathrm{Ca}}(v_{\mathrm{Ca}} - \bar{u}_i)\phi_{z_i} + I \right]. \end{aligned}$$

In this case the application Φ of Theorem 4.4.2 should read as follows for a model with Dirac mass. For $(u, r) \in L^2(I) \times \mathcal{R}_{\mathrm{K}}$, $\Phi(u, r)$ is given by

$$\frac{1}{C}\frac{1}{N_{\mathrm{K}}} \sum_{i=1}^{N_{\mathrm{K}}} c_{\mathrm{K}}(v_{\mathrm{K}} - \bar{u}_i)\phi_{z_i} \int_0^\infty \mathbb{E}_r \left(1_1(r_s^{\mathrm{K},u}(i)) - \frac{\alpha_{\mathrm{K}}(\bar{u}_i)}{\alpha_{\mathrm{K}}(\bar{u}_i) + \beta_{\mathrm{K}}(\bar{u}_i)} \right) \mathrm{d}s,$$

where, for u held fixed, $r_s^{\mathrm{K},u}(i)$ is a Markov chain on $\{0, 1\}$ with jump rate from 1 to 0 is given by $\beta_{\mathrm{K}}(\bar{u}_i)$ and from 0 to 1 by $\alpha_{\mathrm{K}}(\bar{u}_i)$. Of course, in this case, the law of $(r_s^{\mathrm{K},u}(i), s \geq 0)$ can be fully explicited. After some algebra one obtains that $\Phi(u, r)$ is given by

$$\frac{1}{C}\frac{1}{N_{\mathrm{K}}} \sum_{i=1}^{N_{\mathrm{K}}} c_{\mathrm{K}} \frac{v_{\mathrm{K}} - \bar{u}_i}{\alpha_{\mathrm{K}}(\bar{u}_i) + \beta_{\mathrm{K}}(\bar{u}_i)} \left(1_1(r(i)) - \frac{\alpha_{\mathrm{K}}(\bar{u}_i)}{\alpha_{\mathrm{K}}(\bar{u}_i) + \beta_{\mathrm{K}}(\bar{u}_i)} \right) \phi_{z_i}.$$

Then the diffusion operator $(C^{\mathrm{K}}(u)\phi, \psi)_{L^2(I)}$ is given for $u \in L^2(I)$ by

$$\frac{1}{N_{\mathrm{K}}^2} \sum_{i=1}^{N_{\mathrm{K}}} c_{\mathrm{K}}^2 (v_{\mathrm{K}} - \bar{u}_i)^2 \frac{a_{\mathrm{K}}(\bar{u}_i) b_{\mathrm{K}}(\bar{u}_i)}{(\alpha_{\mathrm{K}}(\bar{u}_i) + \beta_{\mathrm{K}}(\bar{u}_i))^3} \bar{\phi}_i \bar{\psi}_i$$

for $(\phi, \psi) \in L^2(I) \times L^2(I)$. From the above expression, we see that for any $u \in L^2(I)$, C^{K} is of trace class in $L^2(I)$. Let us consider, for $t \in [0, T]$

$$Q_t : \phi \in L^2(I) \mapsto \int_0^t e^{\Delta(t-s)} C^{\mathrm{K}}(u_s) e^{\Delta(t-s)} \phi \mathrm{d}s,$$

where $(u_s, s \in [0,T])$ is the averaged limit. From the expression of C^{K} we see that in the $L^2(I)$-sense, Tr Q_t is finite for any $t > 0$.

We present in Figure 4.1 numerical simulations of the slow fast Morris-Lecar model with no Calcium current for various $\varepsilon > 0$. The averaged model (denoted by $\varepsilon = 0$) and the trace of the diffusion operator are also plotted. We set the calcium current equals to zero in our simulations to emphasize the convergence of the slow-fast spatially extended Morris-Lecar model towards the associated averaged model. See [ML81] Figure 2 for simulations of the deterministic finite dimensional Morris-Lecar system with no calcium current. We observe in Figure 4.1 that averaging affects the model in several ways. As ε goes to zero, the averaged number of spikes on a fixed time duration increases until finally form a front wave in the averaged model ($\varepsilon = 0$). In the same time the intensity of the spikes decreases. Let us also mention the fact that the trace of the diffusion operator is higher in the neighborhood of a spike in accordance to [Wai10], Chapter 5, Section 3, where the same phenomenon has been observed for the finite dimensional stochastic Morris-Lecar model.

Appendix 4.A Numerical data for the simulations

Here are the numerical data used for the simulations of the Morris Lecar model

$$\begin{array}{ccc} C = 1, & c_{\mathrm{K}} = 32, & v_{\mathrm{K}} = -70, \\ a = 1, & c_{\mathrm{Ca}} = 0, & v_{\mathrm{Ca}} = 0, \\ R = 0.5, & N_{\mathrm{K}} = 50, & N_{\mathrm{Ca}} = 0. \end{array}$$

The length of the fiber is $l = 0.5$ and the time duration is $T = 2.4$. The impulse I is of the form

$$I(x,t) = \lambda 1_{[0,0.1]}(x)$$

with $\lambda = 300$. The data for the internal resistance R and the capacitance C are arbitraly chosen for the purpose of the simulations. The values for the other parameters correspond to [ML81].

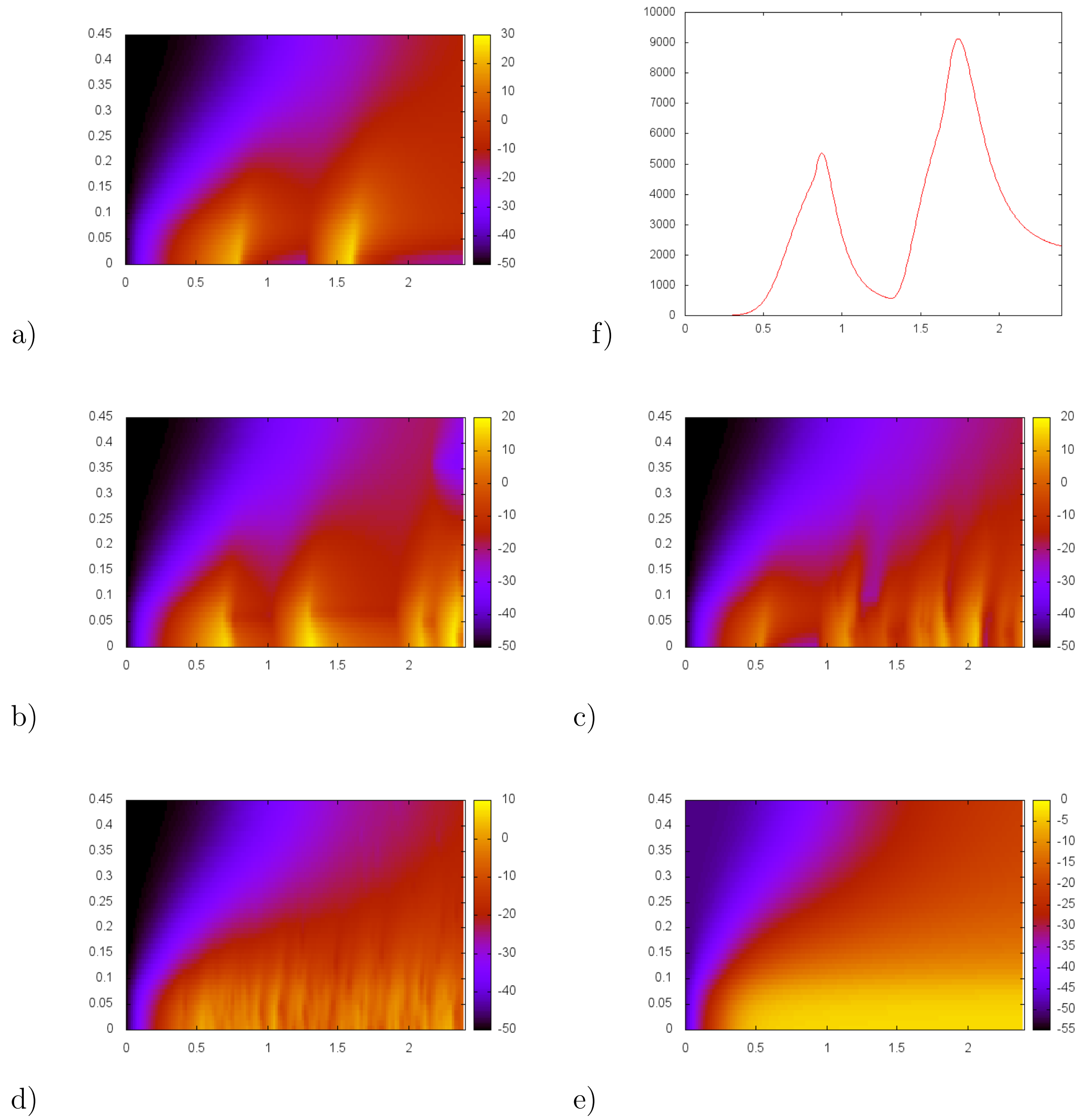


Figure 4.1: Simulations of the spatially extended Morris-Lecar model with no Calcium current for ε equals successively to a) $\varepsilon = 1$, b) $\varepsilon = 0.1$, c) $\varepsilon = 0.01$, d) $\varepsilon = 0.001$, e) $\varepsilon = 0$, that is for the averaged model. The plotted curve f) is related to the simulation of the Morris Lecar model on its left side a): it is the plot of the function $t \mapsto \mathrm{Tr}\ Q_t$. A stimulus is exciting the membrane during all the time duration of the simulation on the portion $[0, 0.1]$ of the fiber.

Chapter 5

On the quantitative ergodicity of infinite dimensional switching systems and application to averaging

The material for Chapter 5 is taken from the submitted pre-publication [GT13b] *On the quantitative ergodicity of infinite dimensional switching systems and application to averaging* available on Arxiv.

5.1 Introduction

This chapter is concerned with the long time behavior of a general class of infinite dimensional Piecewise Deterministic Markov Processes (PDMP). Finite dimensional PDMP have been introduced in [Dav84] and generalized to the infinite dimensional setting in [BR11]. At first, PDMPs were introduced on one hand as a class of general non diffusion processes and on the other hand as tractable models in optimal control theory, see [Dav93], or for more recent works in this field of application, [BDSD12, Gor12]. Then, PDMP systems have been used to model various situation such as motor molecular models [FGC08] and neuron models [BR11]. More generally, PDMPs constitute a very active area of current research, see for example [ADGP12a, BDSD12, FGC08, Gor12, LP13, PTW12, Rie12a, TK09]. Recently, the quantitative and qualitative ergodicity of a class of finite dimensional PDMP have been studied, see [BLBMZ12, BLMZ12]. Our aim in the present chapter is threefold. The first is to extend the results of the paper [BLBMZ12] to cover the case of infinite dimensional PDMP. The second is to propose a class of infinite dimensional PDMP for which the invariant measure exists, is unique and can be made explicit in terms of the invariant measures of finite dimensional PDMPs. At last, the third objective is to apply these results to the averaging of a fast PDMP fully coupled to a slow continuous time Markov chain, a situation motivated by the study of biological systems.

In the present paper, we work with a general class of hybrid systems consisting in an abstract Partial Differential Equation (PDE) coupled to a Continuous Time Markov Chain (CTMC) with finite (in the largest part of the paper) or countable state space. The CTMC is itself coupled to the PDE through its rate of jumps. Therefore, this process decomposes itself in two components: the continuous component solution of the abstract PDE, treated as an Hilbert space valued ordinary differential equation, and the pure jump component following the dynamic of the CTMC. This yields a general class of PDMP as considered in [BR11] and Chapter 3 and 4 for the infinite dimensional setting and [BLBMZ12, BLMZ12] for the finite one. We propose conditions on the linear and non-linear parts of the PDE but also on the jump rates of the CTMC under which the PDMP admits an invariant measure. Then, under some additional assumptions, uniqueness of the invariant measure is obtained. In this case, we show that the convergence of the law of the continuous component toward its invariant measure is exponential in Wasserstein distance. This rate of convergence is obtained by coupling arguments as in [BLBMZ12] but here in the case of *infinite* dimensional PDMP. Finite dimensional projection of our PDMP model are also considered. We show that these truncated processes have respectively unique invariant measures whose first marginals converge, in the sense of the Wasserstein metric, toward the invariant measure of the continuous component of the PDMP. This provide a way to approximate the invariant measure of an Hilbert space valued infinite dimensional PDMP by the use of its finite dimensional projection. These results are then applied to a general class of conductance based neuron model which contains the celebrated Hodgkin-Huxley models, see [Aus08, HH52].

To push forward our analysis, we introduce a class of hybrid model consisting as before in a PDE coupled to a CTMC but with the following changes. On one hand the CTMC is independent of the PDE: the jump rates are no longer coupled to the solution of the PDE. On the other hand we allow the CTMC to take values in a countable set not necessary finite. At last, the PDE has a special structure that we call "diagonal", leading to the study of an infinite system of independent finite dimensional PDMPs driven by independent CTMC, see Section 5.3.1. This diagonal class of PDMP allows us to mimic the situation of the stochastic heat equation, see for example [Wal81], Section 5. In this specific setting, not entirely covered by the results of Section 5.2.1, the rate of convergence toward the unique invariant measure is obtained in Wasserstein distance and the strucure of the invariant measure is made explicit. Convergence of the invariant measures of the finite dimensional projection processes toward the invariant measure of the continuous component is established as well. An example of such a diagonal PDMP for which the invariant measure can be fully computed is then fully treated. This PDMP depends on a parameter β and the long time behavior of the system is analyzed in function of β.

Numerous biological or physical situations where PDMP models are employed exhibit different timescales. See for example [Hil84] in the case of neuron model. Thus, as an application of the result on the rate of convergence in Wasserstein distance of the continuous component of the PDMP towards its invariant measure, we consider the situation of averaging for infinite dimensional PDMP. See [YZ09] for a introduction to averaging methods for inhomogeneous CTMC, a situation different from ours but very instructive. We add a timescale in the general class of hybrid processes studied in this paper. In this framework, the jump part decomposes itself in a slow and a fast component whereas the continuous component is also fast. This ends up with a slow continuous time Markov chain fully coupled to a fast PDMP. This kind of slow-fast systems are common in neurosciences: the nerve impulse as well as some ionic channels have a faster dynamic than some other ionic channels evolving at a slower timescale, see [Hil84] or Section 5.2.2 and Remark 5.4.1 below for more details. We investigate how the system behaves in long time. Using the results mentioned before, we show that the dynamic of the slow component is averaged against the unique invariant measure associated to the fast dynamic. We show that the limiting process is a classical continuous time Markov chain.

The chapter is organized as follows. In Section 5.2.1, we introduce a general class of infinite dimensional PDMPs for which we show the existence and uniqueness of an invariant measure as well as the rate of convergence towards this invariant measure with respect to the Wassertein metric. An example is provided in Section 5.2.2: we show the existence and uniqueness for a general class of conductance based neuron models. In Section 5.3.1, we define a class of "diagonal" PDMPs for which the explicit structure of the invariant measure is obtained. A concrete example is fully treated in Section 5.3.2. As an application of the results of Section 5.2.1, we propose to average a fast PDMP fully coupled to a slow CTMC in Section 5.4.

5.2 A general class of infinite dimensional switching systems

The aim of the section is to extend the results of [BLBMZ12] about the long time behavior of a general class of finite dimensional PDMP models to the infinite dimensional setting. Approximating results are also obtained through the consideration of finite dimensional projections of the model.

5.2.1 Model and results

Let E be a finite set and H a separable Hilbert space. Scalar product and corresponding norm on H are respectively denoted by $(\cdot,\cdot)$ and $\|\cdot\|$. Then, we are led to consider the following family of non-linear problem. For any fixed $i \in E$

$$\begin{cases} \partial_t u_t &= A_i u_t + F_i(u_t), \\ u_0 &\in H \end{cases} \tag{5.1}$$

for $t \in \mathbb{R}_+ = [0, \infty[$. We assume that for each $i \in E$, A_i is a strongly dissipative H-valued linear operator with domain $\mathcal{D}(A_i)$ included in H. This means that there exists a positive constant ω_1^i such that for all $u \in H$

$$(A_i u, u) \leq -\omega_1^i \|u\|^2. \tag{5.2}$$

Since E is a finite set, we can choose ω_1^i independent of i, i.e. set $\omega_1 = \min_{i\in E} \omega_1^i$. For any $i \in E$, the possibly non-linear term F_i is a H-valued operator with domain $\mathcal{D}(F_i)$ included in H, globally Lipschitz continuous on H: there exists a real constant ω_2 such that for any u, v in H and $i \in E$

$$\|F_i(u) - F_i(v)\| \leq \omega_2 \|u - v\|. \tag{5.3}$$

We assume that the evolution problem (5.1) is globally strongly dissipative in the sense that:

$$-\omega_1 + \omega_2 < 0. \tag{5.4}$$

Under the conditions (5.2), (5.3) and (5.4), for any $i \in E$, the system (5.1) has a unique solution u which belongs to $\mathcal{C}([0,T], H)$ for any time horizon $T \geq 0$, see for example [RR04], Chapter 12.

Example 5.2.1. *Let E be a finite set and $H = L^2(0,1)$. For $i \in E$, take $\nu(i)$ and $b(i)$ two positive constants and consider the equation*

$$\partial_t u_t = \nu(i)\Delta u_t - b(i)u + f_i(u) \tag{5.5}$$

on $[0,1]$ with Neumann boundary conditions $\partial_x u_t(0) = \partial_x u_t(1)$. If we assume that f_i is Lipschitz continuous on H with Lipschitz constant ω_2 independent of i, this model enters in our framework if the dissipativity condition $-\min_{i\in E}(\nu(i)+b(i)) + \omega_2 < 0$ is fulfilled.

Going back to the general setting, for any fixed $u \in H$, we consider an E-valued continuous time Markov chain $I = (I_t, t \geq 0)$ with u-dependent generator $Q(u) = (q_{ij}(u))_{i,j}$. We assume that the jump rates are continuous in u and uniformly bounded in u and i from above and below: there exist two constant q_- and q^+ such that

$$0 < q_- < \min_{i \neq j} \inf_u q_{ij}(u) < \max_{i \neq j} \sup_u q_{ij}(u) < q^+ < \infty. \tag{5.6}$$

We propose ourselves to study the following evolution problem

$$\left\{ \begin{array}{rcl} \partial_t u_t &=& A_{I_t} u_t + F_{I_t}(u_t), \\ \mathbb{P}(I_{t+h} = j | I_t = i) &=& q_{ij}(u_t)h + \mathrm{o}(h), \quad i \neq j \end{array} \right. \tag{5.7}$$

for $t \geq 0$ and given initial conditions $u_0 \in H$ and $I_0 \in E$, two random variables. The process $(u_t, I_t)_{t \geq 0}$ is a Piecewise Deterministic Markov Process (PDMP). The existence and uniqueness of a solution to problem (5.7) taking values in $\mathcal{C}([0,T], H) \times \mathbb{D}([0,T], E)$ for any finite time horizon T has been established in [BR11]. Therefore, we are interested in the present paper in the long time behavior of this solution. We proceed by showing that system (5.7) is well defined for any positive time t. Actually, we show that the solution to (5.7) is uniformly bounded. In particular there is no blow up at infinity for the continuous component u. The fact that there is no blow up for the jump component I is a direct consequence of the uniform boundedness (5.6) of the jump rate functions.

Proposition 5.2.1. *Let us assume conditions (5.2), (5.3), (5.4). If the random variable u_0 is bounded in H $\mathbb{P}$-almost surely, so is the process $(u_t, t \geq 0)$: there exists a constant C such that*

$$\sup_{t \in \mathbb{R}_+} \|u_t\| \leq C \quad \mathbb{P} - a.s.$$

Moreover, if

$$r = \frac{\max_{i \in E} \|F_i(0)\|}{(\omega_1 - \omega_2)},$$

the process $(u_t, t \geq 0)$ can not escape from the closed ball of radius r centered at the origin.

Proof. Using successively the dissipativity property of the system and the elementary inequality

$$ab \leq \varepsilon a^2 + \frac{1}{4\varepsilon} b^2$$

which is valid for any positive reals a, b, ε, we see that for all $t \in \mathbb{R}_+$

$$\begin{array}{rcl} \frac{\mathrm{d}}{\mathrm{d}t} \|u_t\|^2 &=& 2(\partial_t u_t, u_t) \\ &=& 2(A_{I_t} u_t + F_{I_t}(u_t), u_t) \\ &\leq& -2\omega_1 \|u_t\|^2 + 2\omega_2 \|u_t\|^2 + 2 \max_{i \in E} \|F_i(0)\| \|u_t\| \\ &\leq& -2(\omega_1 - \omega_2 - \varepsilon) \|u_t\|^2 + 2 \frac{1}{4\varepsilon} \max_{i \in E} \|F_i(0)\|^2 \end{array}$$

$\mathbb{P}$-a.s. Let us define

$$A(\varepsilon) = \omega_1 - \omega_2 - \varepsilon, \quad B(\varepsilon) = \frac{1}{4\varepsilon} \max_{i \in E} \|F_i(0)\|^2.$$

By the Gronwall's lemma we deduce that for all $t \in \mathbb{R}_+$

$$\|u_t\|^2 \leq \frac{B(\varepsilon)}{A(\varepsilon)} \left(1 - e^{-2A(\varepsilon)t}\right) + \|u_0\|^2 e^{-2A(\varepsilon)t}$$

almost-surely. For ε small enough, $A(\varepsilon) > 0$ and this yields

$$\sup_{t \in \mathbb{R}_+} \|u_t\|^2 \leq \max\left(\|u_0\|^2, \frac{B(\varepsilon)}{A(\varepsilon)}\right).$$

For $\varepsilon^* = \frac{\omega_1 - \omega_2}{2}$ we obtain that the stochastic process $(u_t, t \in \mathbb{R}_+)$ can not escape the closed ball of radius

$$r = \sqrt{\left(\frac{\varepsilon^*}{2}\right)^{-1} B\left(\frac{\varepsilon^*}{2}\right)} = \frac{\max_{i \in E} \|F_i(0)\|}{(\omega_1 - \omega_2)}.$$

□

Suppose moreover that there exists a Hilbert basis $\{e_k, k \geq 1\}$ of H which diagonalizes every linear operator A_i and therefore the basis does not depend on i. However the eigenvalues denoted by $-\lambda_k(i)$ may depend on i and we assume that

(C1) For any $i \in E$ the sequence $(\lambda_k(i), k \geq 1)$ is a strictly increasing sequence of positive numbers.

(C2) The following series

$$\sum_{k \geq 1} \frac{1}{(\min_{i \in E} \lambda_k(i))^2}$$

is convergent in $\mathbb{R}$.

As a consequence, the exponentiation of each operator A_i is well defined and

$$\lambda_- = \inf_{k \geq 1} \min_{i \in E} \lambda_k(i) > 0.$$

Under assumptions (C1) and (C2), the process $(u_t, t \in \mathbb{R}_+)$ admits the following mild representation

$$u_t = \sum_{k \geq 1} e^{-\int_0^t \lambda_k(I_s)\mathrm{d}s} (u_0, e_k) e_k + \sum_{k \geq 1} \int_0^t e^{-\int_s^t \lambda_k(I_\tau)\mathrm{d}\tau} (F_{I_s}(u_s), e_k) \mathrm{d}s e_k \tag{5.8}$$

for any $t \in \mathbb{R}_+$, $\mathbb{P}$-a.s.

Example (5.2.1 continued)**.** *These conditions are satisfied in Example 5.2.1. Indeed, for $i \in E$, the linear operator A_i can be defined as*

$$A_i u = \nu(i)\Delta u - b(i)u.$$

If we set $e_1 = 1$ and $e_k = \sqrt{2}\cos((k-1)\pi\cdot)$ for $k \geq 2$, then $\{e_k, k \geq 1\}$ is a Hilbert basis of $L^2(0,1)$ which diagonalizes the linear operator A_i with eigenvalues

$$-\lambda_1 = -b(i), \quad -\lambda_k = -\nu(i)((k-1)\pi)^2 - b(i), \ k \geq 2.$$

These eigenvalues satisfy the two conditions (C1) and (C2).

Proposition 5.2.2. *The family $\{u_t, t \geq 1\}$ is tight in the Hilbert space H.*

Proof. We already know by Proposition 5.2.1 that the family $\{u_t, t \geq 1\}$ is bounded in H almost surely. According to the criteria of tightness in a Hilbert space (cf. Chapter 2) Section 2.3, it remains to show that the family $\{u_t, t \geq 1\}$ can be approximated uniformly in t by a sequence of finite dimensional random variables. Let us define for $N \geq 1$ the truncation of u_t up to the order N as follows

$$u_t^{(N)} = \sum_{k=1}^{N} e^{-\int_0^t \lambda_k(I_s)\mathrm{d}s}(u_0, e_k)e_k + \sum_{k=1}^{N} \int_0^t e^{-\int_s^t \lambda_k(I_\tau)\mathrm{d}\tau}(F_{I_s}(u_s), e_k)\mathrm{d}s e_k$$

for any $t \geq 0$. Then the random variable $u^{(N)}$ lies in the finite dimensional state space $H^{(N)} = \mathrm{span}\{e_k, 1 \leq k \leq N\}$. Moreover, $H^{(N)}$ is an approximation of H in the sense that $\lim_{N\to\infty} \mathrm{d}(w, H^{(N)}) = 0$ for any $w \in H$ where $\mathrm{d}(w, H^{(N)}) = \inf_{v\in H^{(N)}} \|w - v\|$. We are going to show that

$$\lim_{N\to\infty} \sup_{t\geq 1} \|u_t - u_t^{(N)}\| = 0, \quad \mathbb{P} - a.s.$$

For any $t \geq 1$ and $N \geq 1$, using successively the elementary facts that $(a+b)^2 \leq 2a^2 + 2b^2$ and $\min_{i\in E} \lambda_k(i) \leq \lambda_k(j)$ for any $j \in E$, we have

$$\begin{aligned}
&\|u_t - u_t^{(N)}\|^2 \\
= \ & \sum_{k=N+1}^{\infty} \left[e^{-\int_0^t \lambda_k(I_s)\mathrm{d}s}(u_0, e_k) + \int_0^t e^{-\int_s^t \lambda_k(I_\tau)\mathrm{d}\tau}(F_{I_s}(u_s), e_k)\mathrm{d}s \right]^2 \\
\leq \ & 2 \sum_{k=N+1}^{\infty} e^{-2\int_0^t \lambda_k(I_s)\mathrm{d}s}(u_0, e_k)^2 + 2 \sum_{k=N+1}^{\infty} \left(\int_0^t e^{-\int_s^t \lambda_k(I_\tau)\mathrm{d}\tau}(F_{I_s}(u_s), e_k)\mathrm{d}s \right)^2 \\
\leq \ & 2\|u_0\|^2 \sum_{k=N+1}^{\infty} e^{-2\min_{i\in E}\lambda_k(i)t} + 2 \sup_{s\in\mathbb{R}_+} \max_{i\in E} \|F_i(u_s)\|^2 \sum_{k=N+1}^{\infty} \frac{1}{\min_{i\in E} \lambda_k(i)^2},
\end{aligned}$$

$\mathbb{P}$-a.s. The Lipschitz continuity of each F_i, the uniform boundedness of the family $\{u_t, t \in \mathbb{R}_+\}$ proved in Proposition 5.2.1 and the convergence of each of the series of the right hand side independently of $t \geq 1$ yield the desired result. □

According to the assumptions of finiteness of the set E and uniform boundedness of the jump rates of the process $(I_t, t \geq 0)$ we deduce without difficulty the tightness of the process $(I_t, t \geq 0)$ in E. Then the couple process (u, I) is tight in the product space $H \times E$. According to the Prohorov theorem, by tightness, there exists an accumulation point (u^*, I^*) such that

$$\lim_{n\to\infty} (u_{t_n}, I_{t_n}) = (u^*, I^*)$$

in law in $H \times E$ for a certain strictly increasing sequence of real numbers $(t_n, n \geq 1)$. By the Feller property of the process (u, I) and the Krylov-Bogoliubov theorem (see [DPZ92], Chapter 3) we deduce that the law of (u^*, I^*) is invariant for the process $(u_t, I_t, t \geq 0)$.

We are going to show the uniqueness of the invariant measure of the continuous component $(u_t, t \geq 0)$ as well as the rate of convergence of the law of u_t toward the law of u^*. For this purpose, we first recall the definition of the Wasserstein metric.

Definition 5.2.1. *Let ν_1 and ν_2 be two probability laws on a separable and complete metric space $\mathcal{X}$ with finite moments of order $p \in \mathbb{N}$. The Wasserstein distance of order p between ν_1 and ν_2 is defined by*

$$\mathcal{W}_p(\nu_1, \nu_2) = (\inf \mathbb{E}(\|X - Y\|^p_{\mathcal{X}}))^{\frac{1}{p}}\,, \tag{5.9}$$

where the infimum runs over all the couplings (X, Y) of probability laws on $\mathcal{X} \times \mathcal{X}$ with marginals ν_1 and ν_2.

The Wasserstein metric is a metric on the set $\mathcal{P}(\mathcal{X})$ of probabilities on $\mathcal{X}$. $(\mathcal{P}(\mathcal{X}), \mathcal{W}_p)$ is a separable and complete metric space and convergence in Wasserstein distance is equivalent to convergence in law *plus* convergence of the moments of order $1 \leq q \leq p$. When $p = 1$ we simply write $\mathcal{W}$ for $\mathcal{W}_1$. In the present paper, we will successively use the Wasserstein metric of order one and two on the spaces H and $\mathbb{R}$.

For the proof of the following theorem we assume that there exists a constant L_q such that for any $i \in E$ and $u, v \in H$

$$|q_i(u) - q_i(v)| \leq L_q \|u - v\|,$$

where $q_i(u) = \sum_{j\neq i} q_{ij}(u)$ is the total rate for leaving the state i. That is, the total rate functions q_i are Lipschitz in u uniformly in i.

Theorem 5.2.1. *Let (u, I) and $(\tilde{u}, \tilde{I})$ be two solutions of the evolution problem (5.7) such that the support of the laws of the random variables u_0 and $\tilde{u}_0$ are contained in $B[0, r]$. Then*

$$\mathcal{W}(\mathcal{L}(u_t), \mathcal{L}(\tilde{u}_t)) \leq \alpha(1 + t)e^{-\beta t},$$

where the positive constants α and β are made explicit in the proof.

The proof of Theorem 5.2.1 is postponed to Appendix 5.A where we show that the arguments used in the finite dimensional case [BLBMZ12] still work in the infinite dimensional setting. The constants α and β depend on the characteristics of the model, that is on the constants of dissipativity $\omega_1 - \omega_2$, the bound r of the continuous part of the process, the Lipschitz constant and the lower bound of the jump rate functions of the discontinuous part of the process.

As a direct consequence of Theorem 5.2.1, the process u has a unique invariant measure with the following property.

Proposition 5.2.3. *The process $(u_t, t \geq 0)$ has a unique invariant measure ν on H such that*

$$\mathcal{W}(\mathcal{L}(u_t), \nu) \leq \alpha(1+t)e^{-\beta t}$$

where α and β are the constants of Theorem 5.2.1.

Moreover, this invariant measure ν can be approximated in the following way.

Proposition 5.2.4. *For $N \in \mathbb{N}$, let $u^{(N)}$ be the truncation up to the order N of u: $u_t^{(N)} = \sum_{k=1}^N (u_t, e_k)e_k$ for $t \geq 0$. Then $u^{(N)}$ has a unique invariant measure $\nu^{(N)}$ which converges toward ν when N goes to infinity. More precisely, if $\mathbb{E}(\|u_0\|^2)$ is finite we have*

$$\mathcal{W}(\nu^{(N)}, \nu) \leq \sqrt{a_N}$$

for any $N \in \mathbb{N}$ where the sequence $(a_N)_{N\in\mathbb{N}}$ goes to zero when N goes to infinity and is given by

$$a_N = 4(\max_{i\in E} \|F_i(0)\|^2 + C^2) \sum_{k=N+1}^{\infty} \frac{1}{\min_{i\in E} \lambda_k(i)^2}.$$

Proof. The process $u^{(N)}$ has a unique invariant measure $\nu^{(N)}$ on

$$H^{(N)} = \text{span}\{e_k, 1 \leq k \leq N\}$$

by the same arguments as those developed to show the existence and uniqueness of an invariant measure for u or by [BLBMZ12] since $(u^{(N)}, I)$ is a finite dimensional PDMP. To prove the convergence of $\nu^{(N)}$ towards ν we notice that for any $N \in \mathbb{N}$ and $\tau \geq 1$

$$\begin{aligned} &\sup_{t\geq\tau} \|u_t - u_t^{(N)}\|^2 \\ \leq\ & 2\mathbb{E}(\|u_0\|^2) \sum_{k=N+1}^{\infty} e^{-2\min_{i\in E}\lambda_k(i)\tau} + 4(\max_{i\in E} \|F_i(0)\|^2 + C^2) \sum_{k=N+1}^{\infty} \frac{1}{\min_{i\in E} \lambda_k(i)^2} \end{aligned}$$

$\mathbb{P}$-a.s where C is the constant of Theorem 5.2.1, inequality which is proved in Proposition 5.2.2. From this fact, we deduce that for any bounded Lipschitz function f on H with Lipschitz constant less than 1 we have

$$\begin{aligned}&\left|\int f \mathrm{d}\nu^{(N)} - \int f \mathrm{d}\nu\right|^2 \\ \leq\ & 2\mathbb{E}(\|u_0\|^2) \sum_{k=N+1}^{\infty} e^{-2\min_{i\in E}\lambda_k(i)\tau} + 4(\max_{i\in E}\|F_i(0)\|^2 + C^2) \sum_{k=N+1}^{\infty} \frac{1}{\min_{i\in E}\lambda_k(i)^2}.\end{aligned}$$

Then, taking the limit when τ goes to infinity, by dominated convergence and Assumption (C2) we obtain

$$\left|\int f \mathrm{d}\nu^{(N)} - \int f \mathrm{d}\nu\right|^2 \leq 4(\max_{i\in E}\|F_i(0)\|^2 + C^2) \sum_{k=N+1}^{\infty} \frac{1}{\min_{i\in E}\lambda_k(i)^2}.$$

The result follows by the Kantorovich-Rubinstein dual representation of the Wasserstein distance. □

One may want to relax the Lipschitz assumption (5.3) to cover some other interesting situations such as the FitzHugh-Nagumo evolution equation which has been the object of numerous investigation, see for example [San02] and references therein, and is a paradigm in neuroscience since the seminal work [Fit69].

Theorem 5.2.2. *Assume that for $i \in E$ the non linear operator F_i satisfies instead of (5.3) the weaker assumption*

$$(F(u) - F(v), u - v) \leq \tilde{\omega}_2\|u - v\|^2, \tag{5.10}$$

with $\tilde{\omega}_2 \in \mathbb{R}$ such that the condition of global dissipativity $-\omega_1 + \tilde{\omega}_2 < 0$ is satisfied. Then the process $(u_t, t \in \mathbb{R}_+)$ has a unique invariant measure ν which satisfies the estimate of Theorem 5.2.1

$$\mathcal{W}(\mathcal{L}(u_t), \nu) \leq \alpha(1 + t)e^{-\beta t}$$

for α and β two positive constants.

Proof. The proof that u is bounded is identical to the proof of Proposition 5.2.1. The tightness of the process $(u_t, t \in \mathbb{R}_+)$ may present some issues since under assumption (5.10), the boundedness of the family $\{u_t, t \in \mathbb{R}_+\}$ in H is not enough to ensure the boundedness of the family $\{F_i(u_t), t \in \mathbb{R}_+\}$ in H for $i \in E$. However, the estimate

$$\mathcal{W}(\mathcal{L}(u_t), \mathcal{L}(\tilde{u}_t)) \leq \alpha(1 + t)e^{-\beta t}$$

for α and β two positive constants can still be obtained as in the proof of Theorem 5.2.1 and the result follows. □

Example 5.2.2. *Let E be a finite set and $H = L^2(0,1)$. For $i \in E$, let us consider the FitHugh-Nagumo equation*

$$\partial_t u_t = \nu(i)\Delta u_t + b(i)u_t(1-u_t)(u_t-a) \tag{5.11}$$

on $[0,1]$ with Neumann boundary conditions $\partial_x u_t(0) = \partial_x u_t(1)$. Moreover $\nu(i)$ and $b(i)$ are positive constants and $a \in \left(0, \frac{1}{2}\right)$. This model enters in the framework of Theorem 5.2.2 if the dissipativity condition $-\min_{i\in E}(\nu(i)) + \max_{i\in E} b(i) f'\left(\frac{1+a}{3}\right) < 0$ is fulfilled where $f : x \mapsto x(1-x)(x-a)$. Therefore, Theorem 5.2.2 applies and the FitzHugh-Nagumo system (5.11) fully coupled to a CTMC in the sense of (5.7)

$$\left\{\begin{array}{rcl} \partial_t u_t &=& \nu(I_t)\Delta u_t + b(I_t)u_t(1-u_t)(u_t-a) \\ \mathbb{P}(I_{t+h} = j | I_t = i) &=& q_{ij}(u_t)h + \mathrm{o}(h), \quad i \neq j \end{array}\right.$$

has a unique invariant measure ν with exponential convergence in Wasserstein distance of the law of u_t towards the first marginal of ν.

5.2.2 Example: existence and uniqueness of an invariant measure for conductance based neuron models

We are looking for the existence and uniqueness of an invariant measure for the generalized stochastic spatial Hodgkin-Huxley model studied in [Aus08, BR11]. In this model, $H = L^2(0,1)$ is endowed with its usual scalar product and associated norm. The evolution problem is

$$\left\{\begin{array}{rcl} \partial_t u_t &=& \nu\Delta u_t + \frac{1}{N}\sum_{n\in\mathcal{N}} c_{I_t(n)}(v_{I_t(n)} - (u_t, \phi_{z_n}))\phi_{z_n} \\ \mathbb{P}(I_{t+h}(n) = \zeta | I_t(n) = \xi) &=& q_{\xi\zeta}(u_t(z_n))h + \mathrm{o}(h), \quad \xi \neq \zeta. \end{array}\right. \tag{5.12}$$

We still denote by u_0 and I_0 the two initial conditions. The PDE is endowed with zero Dirichlet boundary conditions. The variable u describes the evolution of the action potential along a nerve fiber which is here assimilated to the segment $[0,1]$. The action potential is propagated thanks to the diffusion operator Δ with a certain intensity $\nu > 0$. $\mathcal{N}$ is a finite set of cardinal N. Along the nerve fiber, at location z_n for $n \in \mathcal{N}$, a ion channel is located and can be in a finite number of state denoted by $I(n) \in \mathcal{E}$, where $\mathcal{E}$ is a finite set. Basically a state is to be open or closed. When the ion channel is open, a current is allowed to pass through the membrane. The associated conductance and driven potential for this current are $c_{I(n)} > 0$ and $v_{I(n)} \in \mathbb{R}$. The positive function ϕ_{z_n} is an approximation of the Dirac distribution δ_{z_n}. ϕ_{z_n} says how the opening of a channel affects the local potential of the membrane. For a specific channel at location z_n, if the value of the local potential $u(z_n)$ was fixed equal to $y \in \mathbb{R}$, then the evolution of the states of the ion channels $(I_t(n), t \geq 0)$ would follow a continuous time Markov chain with y-dependent generator $Q(y) = (q_{\xi\zeta}(y))_{\xi\zeta}$. However, the local potential is not fixed. We assume here as in the previous section that the jump rates are bounded above and below uniformly in $y \in \mathbb{R}$.

Remark 5.2.1. *We recall the corresponding deterministic conductance based model. Assume that $z_n = \frac{i}{N}$ for $i \in \{1, \cdots N-1\}$ and that the function ϕ_{z_n} are replaced by the Dirac mass δ_{z_n}. When N goes to infinity, the generalized Hodgkin-Huxley model (5.12) converges in an appropriate sense [Aus08] towards $(v,p) \in \mathcal{C}([0,T], H_0^1(I)) \times \mathcal{C}([0,T], L^2(I)^{|\mathcal{E}|})$ solution of the deterministic generalized Hodgkin-Huxley model*

$$\left\{ \begin{array}{rcl} \partial_t v_t &=& \nu\Delta v_t + \sum_{\xi\in\mathcal{E}} p_{\xi,t} c_\xi (v_\xi - v_t) \\ \partial p_{\xi,t} &=& Q^T(v_t) p_{\xi,t}, \quad \xi \in \mathcal{E}. \end{array} \right. \tag{5.13}$$

Let us write

$$I = (I(n), n \in \mathcal{N}) \in \mathcal{E}^N,$$

which will be referred as a configuration for the ionic channels. Using the formalism of the previous section with $E = \mathcal{E}^N$, the partial differential equation of the system (5.12) becomes:

$$\partial_t u_t = A_{I_t} u_t + F_{I_t}(u_t),$$

where for a given configuration $i \in E$ and a given element $u \in H$ the linear part of the evolution equation is given by

$$A_i u = \nu \Delta u.$$

Notice that for this choice, the operator A_i is independent of i. The reaction term is then defined by

$$F_i(u) = \frac{1}{N} \sum_{n\in\mathcal{N}} c_{i(n)} (v_{i(n)} - (u, \phi_{z_n})) \phi_{z_n}.$$

The following lemma is not difficult.

Lemma 5.2.1. *For any $i \in E$, the linear operator A_i is a self-adjoint operator of negative type. Moreover, if we define, for $k \geq 1$ and $x \in [0,1]$*

$$e_k(x) = \sqrt{2} \sin(k\pi x),$$

then the family $\{e_k, k \geq 1\}$ is a Hilbert basis of $H = L^2(0,1)$ which diagonalizes every operator A_i. The eigenvalues are given by

$$-\lambda_k(i) = -\nu(k\pi)^2.$$

Moreover, the reaction term F_i is dissipative, for any $i \in E$ and $u, v \in H$

$$(F_i(u) - F_i(v), u - v) \leq -\frac{1}{N} \sum_{n\in\mathcal{N}} c_{i(n)} (u - v, \phi_{z_n})^2 \leq 0.$$

We can therefore apply the results of the previous section to obtain

Proposition 5.2.5. *The process $(u_t, t \geq 0)$ has a unique invariant measure μ_u such that if the support of the law of the initial value u_0 is contained in the following closed ball*

$$B\left[0, \frac{\max_{i\in E}\|F_i(0)\|}{\pi}\right],$$

we have

$$\mathcal{W}(\mathcal{L}(u_t), \mu_u) \leq \alpha(1+t)e^{-\beta t}$$

for two positive constants α and β.

5.3 Infinite dimensional switching systems of diagonal type

In this section, we propose a class of infinite dimensional switching systems for which the invariant measure can be described as the convolution of invariant measures of finite dimensional PDMPs. This approach is very similar to the one used when building the invariant measure of the heat equation perturbed by a white noise from the invariant measures of a sequence of finite dimensional Ornstein-Uhlenbeck processes, see Section 5 of [Wal81].

5.3.1 General results

Let $I = (I^k, k \geq 1)$ be a sequence of independent continuous time Markov chains taking values respectively in a finite set E_k, each one being irreducible. As in the previous section, we assume that for each $k \geq 1$, the rate of jumps of I^k are uniformly bounded below and above: there exist two positive constants q_- and q^+ independent of k such that

$$0 < q_- < \min_{i\neq j} q_{ij}^k < \max_{i\neq j} q_{ij}^k < q^+ < \infty. \tag{5.14}$$

The process I is therefore itself an irreducible continuous time Markov chain with values in the countable set $E = \prod_{k\geq 1} E_k$ and such that the transition rates of I are uniformly bounded below and above. We remark here that E can be of infinite cardinality contrary to the previous section. We consider the evolution problem

$$\partial_t u_t = A_{I_t} u_t + F_{I_t}, \tag{5.15}$$

where for $i \in E$ the operators A_i and F_i are diagonal with respect to i. That means that for a Hilbert basis $\{e_k, k \geq 1\}$ of H, on one hand for any $i = (i^k, k \geq 1) \in E$

and $u \in H$

$$A_i u = -\sum_{k\geq 1} \lambda_k(i^k)(u, e_k)e_k,$$

where for any $k \geq 1$ and $i^k \in E_k$, the eigenvalues $\lambda_k(i)$ satisfy Assumptions (C1) and (C2). As before, we assume that the series $\sum(\min_{i\in E_k} \lambda_k(i))^{-2}$ is convergent. On the other hand there exist measurable applications $f_k : E_k \to E_k$ such that

$$F_i = \sum_{k\geq 1} f_k(i^k)e_k$$

with the assumption $\sup_{k\geq 1} \max_{l\in E_k} |f_k(l)| < \infty$. This assumption does not imply that F_i is in H but the semi-group operators associated to A_i regularize the problem such that the process $(u_t, t \geq 1)$ is H-valued. The process $((u_t, I_t), t \in \mathbb{R}_+)$ is a PDMP with constant jump rates. Let us define for $k \geq 1$ and $t \geq 0$, $u_t^k = (u_t, e_k)$. Then $(u_t^k, t \in \mathbb{R}_+)$ is solution of the following finite dimensional evolution problem

$$\frac{\mathrm{d}}{\mathrm{d}t} u_t^k = -\lambda_k(I_t^k)u_t^k + f_k(I_t^k) \tag{5.16}$$

and therefore, (u^k, I^k) is a finite dimensional PDMP. In the following lemma recall that $\mathcal{W}_2$ stands for Wasserstein distance of order 2 on the set of probabilities on $\mathbb{R}$ with finite moment of order two.

Lemma 5.3.1. *For any $k \geq 1$, the PDMP (u^k, I^k) has a unique invariant measure μ^k on $\mathbb{R} \times E^k$ with compact support*

$$C_k = [-r_k, r_k] \times E_k,$$

where $r_k = \frac{\max_{l\in E_k} |f_k(l)|}{\min_{l\in E_k} \lambda_k(l)}$. Moreover, if (u^k, I^k) and $(\tilde{u}^k, \tilde{I}^k)$ are two solutions of (5.16) such that the support of the law of (u_0^k, I_0^k) and $(\tilde{u}_0^k, \tilde{I}_0^k)$ is included in C_k then

$$\mathcal{W}_2^2(\mathcal{L}(u_t^k), \mathcal{L}(\tilde{u}_t^k)) \leq 4r_k^2 \left(e^{-2\min_{l\in E_k} \lambda_k(l)(1-\gamma)t} + e^{-\rho_k \gamma t}\right)$$

for any $\gamma \in (0,1)$ and $\rho_k > 0$ is chosen such that if T_k is the first time of coalescence of I and $\tilde{I}$

$$\mathbb{P}(T^k > t) \leq e^{-\rho_k t}.$$

Such a ρ_k exists by irreducibility of the continuous time Markov chain I^k.

Proof. Let $k \geq 1$ held fixed and (u^k, I^k) be the stochastic process solution of (5.16). We easily see that the process can not escape the compact set $C_k = [-r_k, r_k] \times E^k$ where $r_k = \frac{\max_{i\in E_k} |f_k(i)|}{\min_{i\in E_k} \lambda_k(i)}$. Let (u^k, I^k) and $(\tilde{u}^k, \tilde{I}^k)$ be two solutions of (5.16) with the support of the law of the initial conditions in C_k. Let T_c^k be the first time of coalescence of the two independent processes I^k and $\tilde{I}^k$

$$T_c^k = \inf\{t \geq 0, I_t^k = \tilde{I}_t^k\}.$$

We couple the two processes (u^k, I^k) and $(\tilde{u}^k, \tilde{I}^k)$ in imposing that after the coalescent time T_c, $I_t = \tilde{I}_t$. For any $t \geq 0$ and $\gamma \in (0,1)$ we have

$$\begin{aligned}\mathbb{E}(|u_t^k - \tilde{u}_t^k|^2) &= \mathbb{E}(|u_t^k - \tilde{u}_t^k|^2 1_{\gamma t \geq T_c^k}) + \mathbb{E}(|u_t^k - \tilde{u}_t^k|^2 1_{\gamma t < T_c^k}) \\ &= \mathbb{E}(|u_t^k - \tilde{u}_t^k|^2 1_{\gamma t \geq T_c^k}) + 4r_k^2 \mathbb{P}(\gamma t < T_c^k).\end{aligned}$$

On one hand we know that for the irreducible continuous time Markov chain I^k there exists a positive constant ρ_k such that

$$\mathbb{P}(\gamma t < T_c^k) \leq e^{-\rho_k \gamma t}.$$

On the other hand, on the event $\{\gamma t \geq T_c^k\}$ since $I_t^k = \tilde{I}_t^k$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}|u_t^k - \tilde{u}_t^k|^2 \leq -2 \min_{i \in E_k} \lambda_k(i)(u_t^k - \tilde{u}_t^k)^2.$$

Hence

$$\begin{aligned}|u_t^k - \tilde{u}_t^k|^2 1_{\gamma t \geq T_c^k} &\leq |u_{T_c^k}^k - \tilde{u}_{T_c^k}^k|^2 e^{-2 \min_{i \in E_k} \lambda_k(i)(t - T_c^k)} 1_{\gamma t \geq T_c^k} \\ &\leq 4r_k^2 e^{-2 \min_{i \in E_k} \lambda_k(i) t(1-\gamma)},\end{aligned}$$

almost-surely. Therefore, for any $t \geq 0$ and $\gamma \in (0,1)$ we have

$$\mathbb{E}(|u_t^k - \tilde{u}_t^k|^2) \leq 4r_k^2 \left(e^{-\rho_k \gamma t} + e^{-2 \min_{i \in E_k} \lambda_k(i) t(1-\gamma)}\right)$$

and the result follows. □

We recall that a subset C of $\mathbb{R}^{\mathbb{N}} \times E$ can be viewed as a subset of $H \times E$ by the natural identification between H with $l^2(\mathbb{R})$ and $l^2(\mathbb{R})$ with $\mathbb{R}^{\mathbb{N}}$ where $l^2(\mathbb{R})$ denotes the set of real sequences $(x_k, k \geq 1)$ such that $\sum_{k \geq 1} x_k^2 < \infty$. Since the sequence $(I^k, k \geq 1)$ is a sequence of independent processes, we easily see that

Theorem 5.3.1. *The PDMP (u, I) has a unique invariant measure in $H \times E$ which is of the diagonal form*

$$\mu = \bigotimes_{k \geq 1} \mu^k$$

and supported by the set $C = \prod_{k \geq 1} C_k$.

Recall that by Assumption (C1)

$$\lambda_- = \inf_{k\geq 1} \min_{l\in E_1} \lambda_k(l) > 0$$

and by Assumption (C2), since $\sum_{k\geq 1}(\min_{l\in E_k} |\lambda_k(l)|)^{-2} < \infty$, we have

$$\sum_{k\geq 1} r_k^2 < \infty.$$

We denote by ν the law of the first marginal of μ which is the unique invariant law for the process $(u_t, t \in \mathbb{R}_+)$. Moreover $\nu = \bigotimes_{k\geq 1} \nu^k$ where ν^k is the unique invariant measure of the process $(u_t^k, t \in \mathbb{R}_+)$. We define $\nu^{(N)} = \bigotimes_{k=1}^N \nu^k$ the approximation of ν up to the order N.

Theorem 5.3.2. *Let us assume that all the CTMC I^k have the same law with $E_k = E_1$ for $k \geq 1$. Then*

i) for any $t \in \mathbb{R}_+$, the process $(u_t, t \in \mathbb{R}_+)$ converges towards its invariant measure in the following sense

$$\mathcal{W}_2^2(\mathcal{L}(u_t), \nu) \leq 8\left(\sum_{k\geq 1} r_k^2\right) e^{-\tau t}$$

for any $t \geq 0$ where $\tau = \frac{2\rho_1\lambda_-}{\rho_1+2\lambda_-}$;

ii) for any $N \in \mathbb{N}$, the measure $\nu^{(N)}$ converges towards ν in the following sense

$$\mathcal{W}(\nu^{(N)}, \nu) \leq \sqrt{a_N},$$

where $(a_N)_{N\in\mathbb{N}}$ is the sequence of positive numbers converging to zero given by

$$a_N = 4 \sup_{k\geq 1} \max_{i\in E_1} |f_k(i)|^2 \sum_{k=N+1}^{\infty} \frac{1}{\min_{i\in E} \lambda_k(i)^2}.$$

Proof. i) Let $((u_t, I_t), t \in \mathbb{R}_+)$ and $((\tilde{u}_t, \tilde{I}_t), t \in \mathbb{R}_+)$ be two solutions of problem (5.15). We have, for any $t \geq 0$

$$\mathbb{E}(\|u_t - \tilde{u}_t\|^2) = \sum_{k\geq 1} \mathbb{E}(|u_t^k - \tilde{u}_t^k|^2).$$

For each k, choose a coupling by saying that $I_t^k = \tilde{I}_t^k$ as soon as $t \geq T^k$ where $T^k = \inf\{t \geq 0, I_t^k = \tilde{I}_t^k\}$ (the law of T^k is independent of k because the CTMCs I^k have the same law). By the proof of the previous lemma we have

$$\mathbb{E}(|u_t^k - \tilde{u}_t^k|^2) \leq 4r_k^2 \left(e^{-2\min_{i\in E_1} \lambda_k(i)(1-\gamma)t} + e^{-\rho_1\gamma t}\right)$$

for any $\gamma \in (0,1)$ where ρ_1 is the positive constant such that $\mathbb{P}(T_1 > t) \le \exp(-\rho_1 t)$. Therefore, with $\lambda_- = \inf_{k\ge 1}\min_{l\in E_1}\lambda_k(l)$ we obtain

$$\mathbb{E}(\|u_t - \tilde{u}_t\|^2) = 4\left(e^{-2\lambda_-(1-\gamma)t} + e^{-\rho_1\gamma t}\right)\sum_{k\ge 1} r_k^2.$$

The result follows by taking $\gamma = \frac{2\lambda_-}{\rho_1+2\lambda_-}$.
ii) As in the proof of Proposition 5.2.2, one can show that for any time $\tau \ge 1$

$$\begin{aligned}&\sup_{t\ge\tau}\|u_t^{(N)} - u_t\|^2 \\ &\le 2\mathbb{E}(\|u_0^{(N)} - u_0\|^2)\sum_{k=N+1}^{\infty} e^{-2\min_{i\in E}\lambda_k(i)\tau} + 4\sup_{k\ge 1}\max_{i\in E_1}|f_k(i)|^2\sum_{k=N+1}^{\infty}\frac{1}{\min_{i\in E}\lambda_k(i)^2},\end{aligned}$$

$\mathbb{P}$-a.s. Then the result can be proved as in Proposition 5.2.4. □

5.3.2 Example: a case where the invariant measure is explicit

Let $I = (I^k, k \ge 1)$ be a sequence of i.i.d. continuous time Markov chains with state space $\{0,1\}$ and transition rate matrix

$$Q_\beta = \beta\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

with β a positive parameter. The smaller is β the less the chain jumps. Therefore $E_k = E_1 = \{0,1\}$ for all $k \in \mathbb{N}$. Let g be in $L^2(0,1)$ and $\{e_k, k \ge 0\}$ be the Hilbert basis of $H = L^2(0,1)$ considered in Lemma 5.2.1. We consider the following evolution problem

$$\partial_t u_t = \Delta u_t + \sum_{k\ge 1}(2I_t^k - 1)(g, e_k)e_k \tag{5.17}$$

for $t \ge 0$ with zero Dirichlet boundary conditions. If we freeze the component I to a given $i = (i^k, k \ge 1) \in E$, we are left with the PDE

$$\partial_t u_t = \Delta u_t + \sum_{k\ge 1}(2i^k - 1)(g, e_k)e_k$$

with zero Dirichlet boundary conditions. This PDE gives rise to a very simple dynamical system: there is a unique fixed point v_i which is globally uniformly exponentially stable (GUES)

$$v_i = -\Delta^{-1}\sum_{k\ge 1}(2i^k - 1)(g, e_k)e_k = \sum_{k\ge 1}\frac{(2i^k - 1)}{(k\pi)^2}(g, e_k)e_k.$$

Thus, when I is allowed to switch, the process $(u_t, t \geq 0)$ is successively attracted by the fixed points $\{v_i, i \in E\}$.

Lemma 5.3.2. *For any $k \geq 1$ and $t \geq 0$, the process $u_t^k = (u_t, e_k) \in \mathbb{R}$ is solution to the following problem*

$$u_t^k = -(k\pi)^2 u_t^k + (2I_t^k - 1)(g, e_k).$$

Each process $(u_t^k, I_t^k, t \geq 0)$ has a unique invariant measure μ_β^k supported by the compact set

$$D_k \times E_1,$$

where

$$D_k = \left[-\frac{|(g, e_k)|}{(k\pi)^2}, \frac{|(g, e_k)|}{(k\pi)^2} \right].$$

Moreover the invariant measure is fully explicit: for $x \in \mathbb{R}$ and $i \in \{0, 1\}$, it is given by

$$\begin{aligned} &\mu_\beta^k(\mathrm{d}x, i) \\ = \quad & a_k \left(\frac{1_0(i)}{(g, e_k) - (k\pi)^2 x} + \frac{1_1(i)}{(g, e_k) + (k\pi)^2 x} \right) ((g, e_k)^2 - (k\pi)^4 x^2)^{\frac{\beta}{(k\pi)^2}} 1_{D_k}(x) \mathrm{d}x \end{aligned}$$

where a_k is the constant of normalization defined by

$$(a_k)^{-1} = 2(g, e_k) \int_{D_k} ((g, e_k)^2 - (k\pi)^4 x^2)^{\frac{\beta}{(k\pi)^2} - 1} \mathrm{d}x.$$

By independence of the coordinates, the law of the unique invariant measure of the PDMP $(u_t, I_t, t \geq 0)$ is given by the product of convolution

$$\mu_\beta = \bigotimes_{k \geq 1} \mu_\beta^k.$$

Take $(u^{*,k}, I^{*,k})$ with law μ_β^k. Then the law of $u^{*,k}$ has the density

$$\nu_\beta^k(\mathrm{d}x) = 2(g, e_k) a_k ((g, e_k)^2 - (k\pi)^4 x^2)^{\frac{\beta}{(k\pi)^2} - 1} 1_{D_k}(x) \mathrm{d}x.$$

Then if (u^*, I^*) has the same law as μ_β, we obtain that u^* has the same law as the H-valued random variable

$$Z_\beta = \sum_{k \geq 1} Z_k e_k,$$

where $(Z_k, k \geq 1)$ is a sequence of independent real valued random variables with respective laws ν_β^k. This provides a representation for ν_β, the law of the first marginal of the invariant law μ_β. Let us remark a few things about the random variable Z_β. Firstly, the law ν_β^k for $k \geq 1$ is centered and symmetric such that

$$Z_k = -Z_k \quad \text{in law.}$$

Therefore, we obtain

$$\begin{aligned} Z_\beta &= -Z_\beta & \text{in law,} \\ Z_\beta\left(\frac{1}{2}+x\right) &= Z_\beta\left(\frac{1}{2}-x\right) & \text{in law} \end{aligned}$$

for any $x \in \left[0, \frac{1}{2}\right)$ by a direct computation using the symmetry and the independence of the $\{Z_k, k \geq 1\}$. One can also show that

$$\mathbb{E}(Z_k^2) \sim \frac{(g, e_k)^2}{(k\pi)^4},$$

which ensures that $Z_\beta \in L^2(\Omega, H)$. We remark also that the support of ν_β does not depends on β. However the simulations of Figure 5.1 shows that the set where the measure ν_β is concentrated is highly dependent on β. When β is small (i.e. the jump process I does not jump often), this set concentrates itself around a closed curve connecting 0 to 1. This closed curve corresponds to the union of the two curves defined thanks to the respective unique fixed points of the PDEs

$$\begin{aligned} \partial_t u &= \Delta u + g, \\ \partial_t u &= \Delta u - g \end{aligned}$$

for $t \geq 0$ on $[0, 1]$ with zero Dirichlet boundary conditions. These two fixed points correspond to the two extreme configurations $i^k = 1$ for all $k \geq 1$ and $i^k = 0$ for all $k \geq 1$ in the model (5.17). The fixed points are simply $\Delta^{-1}g$ and $-\Delta^{-1}g$ which are in fact the primitive of the primitive of g (respectively $-g$) which is null in 0 and 1. This yields for the example of Figure 5.1 where $g : x \in [0, 1] \mapsto x(1-x)$

$$\Delta^{-1}g : x \in [0, 1] \mapsto -\frac{1}{12}x^4 + \frac{1}{6}x^3 - \frac{1}{12}x.$$

When β increases, the random variable Z_β starts to visit all the space between the fixed points $\Delta^{-1}g$ and $-\Delta^{-1}g$. Then when β becomes very big (the jump process I jumps very often), the random variable Z_β seems to concentrate itself around the x-axis. As a conclusion one can say that the support of ν_β is the domain delimited by the two functions $\Delta^{-1}g$ and $-\Delta^{-1}g$. For β near zero, the measure is concentrated on the boundary of the support and when β increases there is a "migration" of this concentration toward the x-axis.

Let us gain a more accurate insight in Problem (5.17) when β goes to infinity. Let T be a fixed time horizon. In this case one can show by averaging methods that the process $(u_t, t \in [0,T])$ converges in law when β goes to infinity toward the process $\bar{u}$ solution of the PDE

$$\partial_t \bar{u}_t = \Delta \bar{u}_t + \bar{F}$$

with zero Dirichlet boundary conditions. The averaged reaction term $\bar{F}$ is given by

$$\bar{F} = \int_{\{0,1\}^{\mathbb{N}}} \sum_{k\geq 1} (2i^k - 1)(g, e_k) \bigotimes_{k\geq 1} \pi_k(\mathrm{d}i^k) e_k,$$

where for $k \geq 1$, π_k is the invariant measure of the chain I^k, that is to say $\pi_k(0) = \pi_k(1) = \frac{1}{2}$. Therefore $F = 0$ and

$$\bar{u}_t = e^{\Delta t} u_0$$

for any $t \in [0,T]$. Since we know that $\bar{u}_t$ converges to 0 in H when t goes to infinity, this heuristically explains the behavior of Z when β increases.

If we make, at least formally, the change of time $t \to \beta t$ in (5.17), we end up with the system

$$\beta \partial_t u_t = \Delta u_t + F_{I_t}$$

for $t \geq 0$ where the process $(I_t, t \geq 0)$ jumps now at rate 1. Then, letting formally β go to 0, the process u should verify the equation $\Delta u_t + F_{I_t} = 0$ for $t \geq 0$ and therefore $u_t = \Delta^{-1} F_{I_t}$. This explains heuristically the behavior of the random variable Z when β goes to zero.

5.4 Application: Averaging for a slow continuous time Markov chain fully coupled to a fast infinite dimensional PDMP

In this section, we apply the results of Section 5.2.1, and more particularly Theorem 5.2.1, to averaging.

5.4.1 Model and results

We place ourselves in the setting of Section 5.2.1. For $\varepsilon \in (0,1)$, we consider a PDMP of the following form

$$\left\{\begin{array}{rcl} \partial_t u_t &=& \frac{1}{\varepsilon}\left[A_{I_t^{(1)}, I_t^{(2)}} u_t + F_{I_t^{(1)}, I_t^{(2)}}(u_t)\right], \quad u_0 \in H, \\ \mathbb{P}(I_{t+h}^{(1)} = j | I_t^{(1)} = i) &=& \frac{1}{\varepsilon} q_{ij}^{(1)}(u_t) h + \mathrm{o}(h), \quad i \neq j, \quad i,j \in E^{(1)}, \\ \mathbb{P}(I_{t+h}^{(2)} = j | I_t^{(2)} = i) &=& q_{ij}^{(2)}(u_t) h + \mathrm{o}(h), \quad i \neq j, \quad i,j \in E^{(2)} \end{array}\right. \tag{5.18}$$

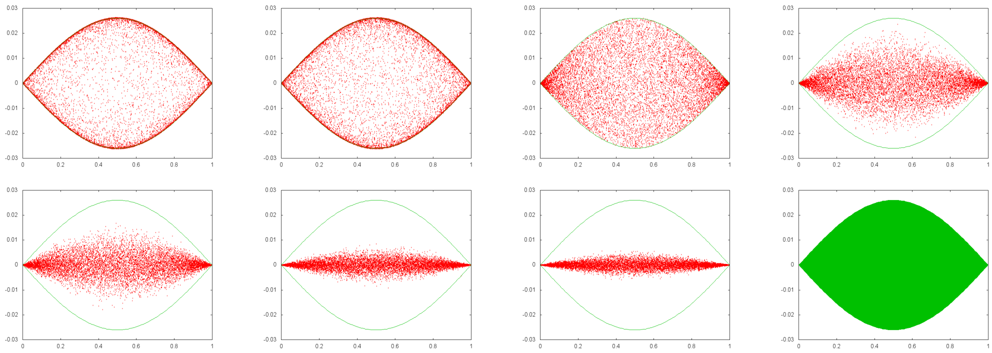


Figure 5.1: Simulation of samples of the H-valued random variable Z for the functions $g : x \in [0,1] \mapsto x(1-x)$ and various values of β: $\beta = 0.1, 1, 10, 50, 100, 500, 1000$ from the upper left corner to the lower right corner. The support of the law is also plotted on the lower right corner. Each point corresponds to a value $Z(x)$ for $x \in [0,1]$.

for $t \geq 0$. Let us describe the above system. If we denote by I the process $(I^{(1)}, I^{(2)})$ then the process (u, I) is nothing but a PDMP as considered in Section 5.2.1. We require on this PDMP the same assumptions as in Section 5.2.1. Particularly

- the sets $E^{(1)}$ and $E^{(2)}$ are finite;
- the operators $A_{(i^{(1)},i^{(2)})}$ and $A_{(i^{(1)},i^{(2)})} + F_{(i^{(1)},i^{(2)})}$ are strongly dissipative on H uniformly in $(i^{(1)}, i^{(2)}) \in E^{(1)} \times E^{(2)}$;
- the jump rates satisfy

  $$0 < q_- < \min_{i \neq j, i,j \in E^{(k)}} \inf_u q_{ij}^{(k)}(u) < \max_{i \neq j, i,j \in E^{(k)}} \sup_u q_{ij}^{(k)}(u) < q^+ < \infty$$

  for $k \in \{1,2\}$ and two positive constants q_- and q^+.
- The jump rate functions $q_{ij} : H \to \mathbb{R}$ are globally Lipschitz uniformly in $i, j \in E^{(k)}$ for $k \in \{1,2\}$.

The PDMP $(u, I^{(1)}, I^{(2)})$ gives rise to two distinct dynamic. The process $(u, I^{(1)})$ evolves faster than the process $I^{(2)}$ according to the timescale separation introduced by the small parameter ε. On a fixed time horizon $[0,T]$, when ε goes to zero, the process $(u, I^{(1)})$ (denoted in the sequel by $(u^\varepsilon, I^{(1),\varepsilon})$) will rapidly reach its stationary behavior. Then the slow process $I^{(2)}$(denoted in the sequel by $I^{2,\varepsilon}$) will evolve according to the averaged dynamic of $(u, I^{(1)})$: the process $(u, I^{(1)})$ will be replaced by its invariant law.

Let us consider that in the system (5.18), the process $I^{(2),\varepsilon}$ is frozen to the value $i^{(2)} \in E^{(2)}$. Then, applying Theorem 5.2.1, we know that the PDMP defined by

$$\begin{cases} \partial_t u_t &= \left[A_{I_t^{(1)},i^{(2)}} u_t + F_{I_t^{(1)},i^{(2)}}(u_t)\right], \quad u_0 \in H, \\ \mathbb{P}(I_{t+h}^{(1)} = j | I_t^{(1)} = i) &= q_{ij}^{(1)}(u_t)h + \mathrm{o}(h), \quad i \neq j, \quad i,j \in E^{(1)} \end{cases} \tag{5.19}$$

has a unique invariant measure $\mu_{i^{(2)}}$ on $H \times E^{(1)}$ such that if $\nu_{i^{(2)}} = \int_{E^{(1)}} \mu_{i^{(2)}}(du, di)$ is the law of its first marginal, we have

$$\mathcal{W}(\mathcal{L}(u_t), \nu_{i^{(2)}}) \leq \alpha(1+t)e^{-\beta t}, \tag{5.20}$$

where the positive constants α and β are independent of $i^{(2)} \in E^{(2)}$. The fact that α and β can be chosen independently of $i^{(2)}$ is a direct consequence of the explicit form of these constants that can be found at the end of Appendix 5.A. As a result of the Kantorovich-Rubinstein dual representation of the Wasserstein distance, we obtain that for any globally Lipschitz functional Φ on H we have

$$\left| \mathbb{E}(\Phi(u_t)) - \int_H \Phi(u) \nu_{i^{(2)}}(\mathrm{d}u) \right| \leq \|\Phi\|_{\mathrm{Lip}} \alpha(1+t)e^{-\beta t}. \tag{5.21}$$

Let us average the dynamic of $I^{(2)}$ against the invariant measure of $(u, I^{(1)})$. For any $i, j \in E^{(2)}$ we define the averaged jump rate

$$\bar{q}_{ij} = \int_H q_{ij}^{(2)}(u) \nu_i(\mathrm{d}u) \tag{5.22}$$

We denote by $\bar{Q}$ the intensity matrix associated to the averaged jump rates $\bar{q}_{ij}$ and by $\bar{J} = (\bar{J}_t, t \in [0,T])$ the continuous time Markov chain associated to $\bar{Q}$. For simplicity, its starting point is assumed to be the one of $I^{(2),\varepsilon}$: $\bar{J}_0 = I_0^{(2),\varepsilon} = i_0^{(2)} \in E^{(2)}$.

Theorem 5.4.1. *The process $I^{(2),\varepsilon} = (I_t^{(2),\varepsilon}, t \in [0,T])$ converges in law when ε goes to zero toward the CTMC $\bar{J}$. Moreover the order of convergence is 1 in the sense that*

$$\sup_{t \geq 0} |\mathbb{E}(\phi(I_t^{(2),\varepsilon}) - \phi(\bar{J}_t))| = \mathrm{O}(\varepsilon)$$

for any real valued measurable and bounded function ϕ.

Proof. The proof is postponed to Section 5.4.2. □

Actually, we show in the proof of Theorem 5.4.1 that for any real valued measurable and bounded function ϕ

$$\mathbb{E}(\phi(I_t^{(2),\varepsilon}) - \phi(\bar{J}_t)) = \mathrm{O}\left(\varepsilon + e^{-\beta \frac{t}{\varepsilon}}\right)$$

for any $t \in [0,T]$ where the big O is uniform in $t \in [0,T]$. This result can be considered as a starting point for the study of the rescaled occupation measure

$$\frac{1}{\sqrt{\varepsilon}} \int_0^t \left(1_i(I_s^{(2),\varepsilon}) - 1_i(\bar{J}_s)\right) \gamma_i(s) \mathrm{d}s$$

for $t \in [0,T]$, $i \in E^{(2)}$ and γ_i a deterministic function. See [YZ09], chapter 5, for more details in the case of non homogeneous continuous time Markov chains.

Remark 5.4.1. *We consider the generalized stochastic spatial Hodgkin-Huxley model of Section 5.2.2 with two timescales*

$$\left\{\begin{array}{rcl} \partial_t u_t^\varepsilon &=& \frac{1}{\varepsilon}\left[\nu \Delta u_t^\varepsilon + F^{(1)}(u_t^\varepsilon, I^{(1),\varepsilon}) + F^{(2)}(u_t^\varepsilon, I^{(2),\varepsilon})\right] \\ \mathbb{P}(I_{t+h}^{(1),\varepsilon}(n) = \zeta | I_t^{(1),\varepsilon}(n) = \xi) &=& \frac{1}{\varepsilon} q_{\xi\zeta}^{(1)}((u_t^\varepsilon, \phi_{z_n}))h + \mathrm{o}(h),\ \xi \neq \zeta,\ \xi,\zeta \in \mathcal{E}^{(1)} \\ \mathbb{P}(I_{t+h}^{(2),\varepsilon}(n) = \zeta | I_t^{(2),\varepsilon}(n) = \xi) &=& q_{\xi\zeta}^{(2)}((u_t^\varepsilon, \phi_{z_n}))h + \mathrm{o}(h),\ \xi \neq \zeta,\ \xi,\zeta \in \mathcal{E}^{(2)}, \end{array}\right. \tag{5.23}$$

where for $l = 1,2$ and $(u, I^l) \in H \times \mathcal{E}^{N_l}$ (keeping the notations of Section 5.2.2)

$$F^{(l)}(u, I^{(l)}) = \frac{1}{N_l} \sum_{n \in \mathcal{N}_l} c_{I_t^{(l)}(n)} (v_{I_t^{(l)}(n)} - (u_t, \phi_{z_n}))\phi_{z_n}.$$

That is, we distinguish two different kind of ion channels located at discrete points in the two finite subsets $\mathcal{N}_1$ and $\mathcal{N}_2$ of $[0,1]$. After averaging, this system reduces to the CTMC $\bar{J}$ with rate of jumps given by

$$\bar{q}_{ij} = \int_H q_{ij}^{(2)}(u) \nu_i(\mathrm{d}u) \tag{5.24}$$

for $i,j \in (\mathcal{E}^{(2)})^{N_2}$. The measure ν_i is the invariant measure corresponding to the equation on the potential u when the process $I^{(2),\varepsilon}$ is held fixed as explained above. Notice that the equation on the potential disappear in the averaged model. The potential is only present through the invariant measure ν. This may look odd since the potential is a variable of first interest in conductance based neuron models. As remarked in [RW07] for the finite dimensional deterministic Hodgkin-Huxley model, there is certainly a transition between the three dimensional two timescales model (5.23) and the one dimensional averaged model following the dynamic of $\bar{J}$. This transition should obey the following evolution problem

$$\left\{\begin{array}{rcl} \partial_t u_t^\varepsilon &=& \frac{1}{\varepsilon}\left[\nu \Delta u_t^\varepsilon + \bar{F}^{(1)}(u_t^\varepsilon) + F^{(2)}(u_t^\varepsilon, I^{(2),\varepsilon})\right] \\ \mathbb{P}(I_{t+h}^{(2),\varepsilon}(n) = \zeta | I_t^{(2),\varepsilon}(n) = \xi) &=& q_{\xi\zeta}^{(2)}((u_t, \phi_{z_n}))h + \mathrm{o}(h), \quad \xi \neq \zeta, \quad \xi,\zeta \in \mathcal{E}^{(2)}, \end{array}\right. \tag{5.25}$$

where the averaged reaction term is given by

$$\bar{F}(u) = \int_{(\mathcal{E}^{(1)})^{N_1}} F^{(1)}(u,i) \pi_u(\mathrm{d}i)$$

and the measure π_u is the invariant measure of the process defined by

$$\mathbb{P}(\tilde{I}^{(1)}_{t+h}(n) = \zeta | \tilde{I}^{(1)}_t(n) = \xi) = q^{(1)}_{\xi\zeta}((u, \phi_{z_n}))h + \mathrm{o}(h), \quad \xi \neq \zeta, \quad \xi, \zeta \in \mathcal{E}^{(1)}$$

for a given u (for averaging in this context , see Chapter 2). System (5.25) constitutes a singularly perturbed model in a suitable form for biological considerations. This transition phenomenon certainly deserves a more profound investigation.

5.4.2 Proof of Theorem 5.4.1

Since the family $\{I^{(2),\varepsilon}, \varepsilon \in (0,1)\}$ takes values in a finite state space $E^{(2)}$ with jump rates which satisfy estimates

$$0 < q_- < \min_{i \neq j} \inf_u q^{(2)}_{ij}(u) < \max_{i \neq j} \sup_u q^{(2)}_{ij}(u) < q^+ < \infty,$$

the tightness of the family in $\mathbb{D}(\mathbb{R}_+, E^{(2)})$ follows. Let J be any accumulation point. We want to identify the generator of the process J. We follow a martingale approach. For $\varepsilon \in (0,1)$, the generator of the process $(u^\varepsilon, I^{(1),\varepsilon}, I^{(2),\varepsilon})$ is given by

$$\begin{aligned} \mathcal{A}^\varepsilon \phi(u, i^{(1)}, i^{(2)}) &= \frac{1}{\varepsilon} \frac{\mathrm{d}}{\mathrm{d}u} \phi(u, i^{(1)}, i^{(2)})[A_{i^{(1)},i^{(2)}} u + F_{i^{(1)},i^{(2)}}(u)] && (5.26)\\ &+ \frac{1}{\varepsilon} \sum_{j \in E^1} [\phi(u, j, i^{(2)}) - \phi(u, i^{(1)}, i^{(2)})] q^{(1)}_{i^{(1)} j}(u) && (5.27)\\ &+ \sum_{j \in E^2} [\phi(u, i^{(1)}, j) - \phi(u, i^{(1)}, i^{(2)})] q^{(2)}_{i^{(2)} j}(u) && (5.28) \end{aligned}$$

for function $\phi : H \times E^1 \times E^2 \to \mathbb{R}$, $\mathcal{C}^1$ and bounded for bounded arguments in its first component and measurable and bounded in its second and third components. The usual theory of Markov processes (see [EK86], Chapter 4) tels us that the stochastic process defined by:

$$M^{\varepsilon,\phi}_t = \phi(u^\varepsilon_t, I^{(1),\varepsilon}_t, I^{(2),\varepsilon}_t) - \phi(u_0, i^{(1)}_0, i^{(2)}_0) - \int_0^t \mathcal{A}^\varepsilon \phi(u^\varepsilon_s, I^{(1),\varepsilon}, I^{(2),\varepsilon}_s) \mathrm{d}s$$

is a $\mathcal{F}^\varepsilon_t$-martingale where the σ-algebra $\mathcal{F}^\varepsilon_t$ is defined for $t \geq 0$ by

$$\mathcal{F}^\varepsilon_t = \sigma(u^\varepsilon_s, I^{(1),\varepsilon}_s, I^{(2),\varepsilon}_s, 0 \leq s \leq t).$$

In particular, for $\phi_1 : E^{(2)} \to \mathbb{R}$ a bounded measurable function we have that the process:

$$M^{\varepsilon,\phi_1}_t = \phi_1(I^{(2),\varepsilon}_t) - \phi_1(i^{(2)}_0) - \int_0^t \sum_{i \in E^{(2)}} [\phi_1(i) - \phi_1(I^{(2),\varepsilon}_s)] q^{(2)}_{I^{(2),\varepsilon}_s i}(u^\varepsilon_s) \mathrm{d}s \qquad (5.29)$$

is a $\mathcal{F}^\varepsilon_t$-martingale. Using this fact, we will show that the process:

$$M^{\phi_1}_t = \phi_1(J_t) - \phi_1(i^{(2)}_0) - \int_0^t \sum_{i\in E^{(2)}} [\phi_1(i) - \phi_1(J_s)]\bar{q}_{J_s i}\mathrm{d}s \tag{5.30}$$

is a $\mathcal{F}_t$-martingale. Let $t, \sigma \geq 0$ and $0 \leq t_1 < t_2 < \cdots < t_k \leq t$ a collection of k real numbers. For bounded continuous functions $z_1, \cdots, z_k$ on $E^{(2)}$ we compute:

$$\begin{aligned}
& \mathbb{E}((M^{\phi_1}_{t+\sigma} - M^{\phi_1}_t)z_1(J_{t_1})\cdots z_k(J_{t_k})) \\
= \ & \mathbb{E}((\phi_1(J_{t+\sigma}) - \phi_1(J_t) - \int_t^{t+\sigma} \sum_{i\in E^{(2)}} [\phi_1(i) - \phi_1(J_s)]\bar{q}_{J_s i}\mathrm{d}s)z_1(J_{t_1})\cdots z_k(J_{t_k})) \\
= \ & \lim_{\varepsilon\to 0} \mathbb{E}((\phi_1(I^{(2),\varepsilon}_{t+\sigma}) - \phi_1(I^{(2),\varepsilon}_t) \\
- \ & \int_t^{t+\sigma} \sum_{i\in E^{(2)}} [\phi_1(i) - \phi_1(I^{(2),\varepsilon}_s)]\bar{q}_{I^{(2),\varepsilon}_s i}(v^i)\mathrm{d}s)z_1(I^{(2),\varepsilon}_{t_1})\cdots z_k(I^{(2),\varepsilon}_{t_k}))
\end{aligned}$$

where we have used the convergence in law of the family $\{(I^{(2),\varepsilon}_t)_{t\in[0,T]}, \varepsilon \in]0,1]\}$ towards $(J_t)_{t\in[0,T]}$ when ε goes to zero. Let us write

$$Z^\varepsilon_k = z_1(I^{(2),\varepsilon}_{t_1})\cdots z_k(I^{(2),\varepsilon}_{t_k}).$$

Splitting the sum in two parts, we obtain

$$\begin{aligned}
& \lim_{\varepsilon\to 0} \mathbb{E}((\phi_1(I^{(2),\varepsilon}_{t+\sigma}) - \phi_1(I^{(2),\varepsilon}_t) - \int_t^{t+\sigma} \sum_{i\in E^{(2)}} [\phi_1(i) - \phi_1(I^{(2),\varepsilon}_s)]\bar{q}_{I^{(2),\varepsilon}_s i}\mathrm{d}s)Z^\varepsilon_k) \\
= \ & \lim_{\varepsilon\to 0} A_\varepsilon + B_\varepsilon,
\end{aligned}$$

where

$$A_\varepsilon = \mathbb{E}((\phi_1(I^{(2),\varepsilon}_{t+\sigma}) - \phi_1(I^{(2),\varepsilon}_t) - \int_t^{t+\sigma} \sum_{i\in E^{(2)}} [\phi_1(i) - \phi_1(I^{(2),\varepsilon}_s)]q_{I^{(2),\varepsilon}_s i}(u^\varepsilon_s)\mathrm{d}s)Z^\varepsilon_k),$$

$$B_\varepsilon = \mathbb{E}((\int_t^{t+\sigma} \sum_{i\in E^{(2)}} [\phi_1(i) - \phi_1(I^{(2),\varepsilon}_s)][q_{I^{(2),\varepsilon}_s i}(u^\varepsilon_s) - \bar{q}_{I^{(2),\varepsilon}_s i}]\mathrm{d}s)Z^\varepsilon_k).$$

At this point, we notice that for the first term

$$\lim_{\varepsilon\to 0} A_\varepsilon = \lim_{\varepsilon\to 0} \mathbb{E}((M^{\varepsilon,\phi_1}_{t+\sigma} - M^{\varepsilon,\phi_1}_t)Z^\varepsilon_k) = 0$$

by the martingale property of M^{ε,ϕ_1}. Then, for the second term, rearranging and conditioning by the σ-algebra $\mathcal{G}_s^\varepsilon = \sigma(I_\rho^{(2),\varepsilon}, 0 \le \rho \le s)$ generated by $I^{(2),\varepsilon}$

$$
\begin{aligned}
&\lim_{\varepsilon\to 0} B_\varepsilon \\
= &\lim_{\varepsilon\to 0}\int_t^{t+\sigma} \mathbb{E}(\sum_{i\in E^{(2)}} [\phi_1(i) - \phi_1(I_s^{(2),\varepsilon})][q_{I_s^{(2),\varepsilon}i}(u_s^\varepsilon) - \bar{q}_{I_s^{(2),\varepsilon}i}]Z_k^\varepsilon)\mathrm{d}s \\
= &\lim_{\varepsilon\to 0}\int_t^{t+\sigma} \mathbb{E}(\sum_{i\in E^{(2)}} [\phi_1(i) - \phi_1(I_s^{(2),\varepsilon})]\mathbb{E}(q_{I_s^{(2),\varepsilon}i}(u_s^\varepsilon) - \bar{q}_{I_s^{(2),\varepsilon}i}|\mathcal{G}_s^\varepsilon)Z_k^\varepsilon)\mathrm{d}s.
\end{aligned}
$$

Using the uniform Lipschitz property of q_{ji} on H and the estimate (5.21) we obtain

$$
\mathbb{E}(q_{I_s^{(2),\varepsilon}i}(u_s^\varepsilon) - \bar{q}_{I_s^{(2),\varepsilon}i}|\mathcal{G}_s^\varepsilon) \le \mathbb{E}(L_q\alpha(1+\frac{s}{\varepsilon})e^{-\beta\frac{s}{\varepsilon}}|\mathcal{G}_s^\varepsilon).
$$

Therefore, including the above estimate in the calculation

$$
\begin{aligned}
&|\lim_{\varepsilon\to 0}\int_t^{t+\sigma} \mathbb{E}(\sum_{i\in E^{(2)}} [\phi_1(i) - \phi_1(I_s^{(2),\varepsilon})]\mathbb{E}(q_{I_s^{(2),\varepsilon}i}(u_s^\varepsilon) - \bar{q}_{I_s^{(2),\varepsilon}i}|\mathcal{G}_s^\varepsilon)Z_k^\varepsilon)ds| \\
\le\ &\lim_{\varepsilon\to 0}\int_t^{t+\sigma} \mathbb{E}(\sum_{i\in E^{(2)}} \left|\phi_1(i) - \phi_1(I_s^{(2),\varepsilon})\right| \mathbb{E}(L_q\alpha(1+\frac{s}{\varepsilon})e^{-\beta\frac{s}{\varepsilon}}|\mathcal{G}_s^\varepsilon)\,|Z_k^\varepsilon|)\mathrm{d}s \\
\le\ &C_1\lim_{\varepsilon\to 0}\int_t^{t+\sigma}\left(1+\frac{s}{\varepsilon}\right)e^{-\beta\frac{s}{\varepsilon}}\mathrm{d}s \\
\le\ &C_2\left(\varepsilon + e^{-\beta\frac{t}{\varepsilon}}\right),
\end{aligned}
$$

where C_1 and C_2 are two constants independent of ε, t and σ. Therefore for any $t \in (0, T]$ and $\sigma > 0$:

$$
\mathbb{E}((M_{t+\sigma}^{\phi_1} - M_t^{\phi_1})z_1(I_{t_1}^{(2),\varepsilon})\cdots z_k(I_{t_k}^{(2),\varepsilon})) = 0.
$$

By the same arguments, one shows that the above equality is still valid for $t = 0$. Therefore, M^{ϕ_1} is a $\mathcal{F}_t$-martingale for any bounded measurable function ϕ_1. We have thus identified J as the process with generator:

$$
\mathcal{A}\phi_1(j) = \sum_{i\in E^{(2)}} [\phi_1(i) - \phi_1(j)]\bar{q}_{ji}.
$$

This process is uniquely defined up to indistinguishability if the constants $\bar{q}_{ji}$ are given and thus the accumulation point is unique. Similarly, we can show that:

$$
\sup_{t\ge 0} |\mathbb{E}(\phi_1(I_t^{(2),\varepsilon}) - \phi_1(J_t))| = \mathrm{O}(\varepsilon)
$$

This implies that the order of convergence is 1.

Appendix 5.A Proof of Theorem 5.2.1

Our proof follows the same arguments as those introduced in the finite dimensional case [BLBMZ12]. Given two real valued random variables X and Y, we say that x is stochastically smaller than Y if for all $x \in \mathbb{R}$, $\mathbb{P}(X \leq x) \leq \mathbb{P}(Y \leq x)$. We keep then notation of Theorem 5.2.1.

Let $((u, I), (\tilde{u}, \tilde{I}))$ be a coupling such that the law of u_0 and $\tilde{u}_0$ are supported in $B[0, r]$. $((u_t, I_t), t \in \mathbb{R}_+)$ and $((\tilde{u}_t, \tilde{I}_t), t \in \mathbb{R}_+)$ are thus two solutions of Problem (5.7) . We begin by noticing that for any $t \geq 0$ we have:

$$\begin{aligned}\|u_t - \tilde{u}_t\|^2 \leq \|u_0 - \tilde{u}_0\|^2 - 2(\omega_1 - \omega_2) \int_0^t \|u_s - \tilde{u}_s\|^2 \mathrm{d}s \\ + 2 \int_0^t (F_{I_s}(\tilde{u}_s) - F_{\tilde{I}_s}(\tilde{u}_s), u_s - \tilde{u}_s) 1_{I_s \neq \tilde{I}_s} \mathrm{d}s\end{aligned}$$

almost surely. Let us denote the quantity $\|u_t - \tilde{u}_t\|$ by D_t (the distance in H-norm between u_t and $\tilde{u}_t$). We see that as long as $I = \tilde{I}$, D_t decreases exponentially fast at rate $\omega_1 - \omega_2$. When $I \neq \tilde{I}$, D_t can increase but will nevertheless remain less than $2r$ since $D_t = \|u_t - \tilde{u}_t\| \leq 2r$ (see Proposition 5.2.1). At worst D_t will reach $2r$ before the time of coalescence of the two processes I and $\tilde{I}$. We have to construct our coupling such that we are able to control the coalescent time and also the time during which the two processes I and $\tilde{I}$ remain equal. This can be achieved using the infinitesimal generator A^c of the process $(u, I, \tilde{u}, \tilde{I})$ as detailed below.

1. if $i \neq \tilde{i}$

$$\begin{aligned}A^c h(u, i, \tilde{u}, \tilde{i}) &= \frac{dh}{du}(u, i, \tilde{u}, \tilde{i})[A_i u + F_i(u)] + \frac{dh}{d\tilde{u}}(u, i, \tilde{u}, \tilde{i})[A_{\tilde{i}} \tilde{u} + F_{\tilde{i}}(\tilde{u})] \\ &+ \sum_{k \in E} q_{ik}(u)(h(u, k, \tilde{u}, \tilde{i}) - h(u, i, \tilde{u}, \tilde{i})) \quad \text{a single jump of } I \\ &+ \sum_{k \in E} q_{\tilde{i}k}(\tilde{u})(h(u, i, \tilde{u}, k) - h(u, i, \tilde{u}, \tilde{i})) \quad \text{a single jump of } \tilde{I}\end{aligned}$$

2. if $i = \tilde{i}$

$$
\begin{aligned}
A^c h(u,i,\tilde{u},\tilde{i}) = & \frac{dh}{du}(u,i,\tilde{u},i)[A_i u + F_i(u)] + \frac{dh}{d\tilde{u}}(u,i,\tilde{u},i)[A_i \tilde{u} + F_i(\tilde{u})] \\
& + \sum_{k\in E}(q_{ik}(u) \wedge q_{ik}(\tilde{u}))(h(u,k,\tilde{u},k) - h(u,i,\tilde{u},i)) \\
& \quad \text{two simultaneous jumps ...} \\
& + \sum_{k\in E}(q_{ik}(u) - q_{ik}(\tilde{u}))^+(h(u,k,\tilde{u},i) - h(u,i,\tilde{u},i)) \\
& \text{... a single jump of } I \\
& + \sum_{k\in E}(q_{ik}(u) - q_{ik}(\tilde{u}))^-(h(u,i,\tilde{u},k) - h(u,i,\tilde{u},i)) \\
& \text{... a single jump of } \tilde{I}
\end{aligned}
$$

Remark 5.A.1. *Let us consider two neurons whose dynamic is described by the conductance based model of Section 5.2.2. Then the proposed coupling corresponds to a coupling on the conductances: the two neurons are linked throughout the rate of jumps of their ionic channels and therefore the related conductances are coupled.*

In between two coincidences of the processes I and $\tilde{I}$, D_t can increase up to $2r$ before the coalescent time T_c of I and $\tilde{I}$. For the proposed coupling where the law of $(u, I, \tilde{u}, \tilde{I})$ is characterized by the above generator, the coalescent time T_c is stochastically smaller than an exponential variable of parameter b for some positive b. For instance, if $E = \{1, 2\}$, the coalescence time is equal to the time of the first jump of one of the two independent processes r and $\tilde{r}$. This event occurs at rate $q_{i_1\tilde{i}_1}(u) + q_{i_2\tilde{i}_2}(\tilde{u}) \geq 2q_-$ with $i_k, \tilde{i}_k \in E = \{1, 2\}$ for $k = 1, 2$.

Then, after a coalescent time, the two processes I_t and $\tilde{I}_t$ remain equal up to the first single jump of one of the two processes which occurs, according to our coupling, at rate $|q_i(u)-q_i(\tilde{u})|$ bounded above by $L_q D_t$. Then $\mathbb{E}(D_t) \leq \mathbb{E}(U_t)$ where U_t, in a way, caricatures the dynamic of the process D_t. U_t decreases exponentially fast at rate $\omega_1 - \omega_2$ and can jump from its current state x to $2r + \varepsilon$ with $\varepsilon > 0$ at rate $L_q x$. This caricature the fact that D_t decreases exponentially fast when I and $\tilde{I}$ remain equal and then can increase up to $2r$. Here U_t increases by jumping directly to $2r+\varepsilon$. When U_t is in the state $2r+\varepsilon$ it stays there until it jumps in $2r$ and that at rate b. This caricature the fact that when the processes I and $\tilde{I}$ are different and thus D_t can increase up to $2r$ but will recover a decreasing phase at the time where I and $\tilde{I}$ become equal and this happens after a time (stochastically) smaller than an exponential variable with parameter b. The generator of U is given by:

$$\begin{aligned} Gf(x) &= [-(\omega_1-\omega_2)xf'(x) + L_q x(f(2r+\varepsilon) - f(x))]\, 1_{[0,2r]}(x) \\ &+ [b(f(2r) - f(2r+\varepsilon))]\, 1_{2r+\varepsilon}(x). \end{aligned}$$

Then, by Theorem 3.1. of [BLBMZ12] we have that there exists two constants $\gamma, c > 0$ such that:

$$\mathbb{E}(U_t) \leq 2r(1+ct)e^{-\eta t},$$

where $\eta = \frac{\omega_1-\omega_2}{1+\frac{\omega_1-\omega_2}{\gamma}}$ with

$$\gamma = \frac{(\omega_1-\omega_2+b) - \sqrt{(\omega_1-\omega_2+b)^2 - 4bp(\omega_1-\omega_2)}}{2}$$

and

$$c = \frac{\omega_1-\omega_2}{\omega_1-\omega_2+\gamma}\frac{ep(\omega_1-\omega_2)b}{\sqrt{(\omega_1-\omega_2+b)^2 - 4bp(\omega_1-\omega_2)}}$$

with $e = \exp(1)$ and $p = \exp\left(-\frac{2rL_q}{\omega_1-\omega_2}\right)$. This concludes the proof.

Chapter 6

Simulations of stochastic partial differential equations for excitable media using finite elements

The present chapter is based on a work in collaboration with M. Thieullen and M. Boulakia. M. Boulakia is member of Laboratoire Jacques-Louis Lions of the Université Pierre et Marie Curie. The preprint corresponding to this chapter is [BGT13].

6.1 Introduction

The present chapter is concerned with the numerical simulation of Stochastic Partial Differential Equations (SPDEs) used to model excitable cells in order to analyze the effect of noise on such biological systems. Our aim is twofold. The first is to propose an efficient and easy-to-implement method to simulate this kind of models. We focus our work on practical numerical implementation with software used for deterministic PDEs such as FreeFem++ or equivalent. The second is to analyze the effect of noise on these systems thanks to numerical experiments. Namely, in models for cardiac cells, we investigate the possibility of purely noise induced reentrant patterns such as spiral or scroll-waves as these phenomena are related to major troubles of the cardiac rhythm such as tachyarrhythmia. For numerical experiments, we focus on the Barkley and Mitchell-Schaeffer models, both originally deterministic models to which we add a noise source.

Mathematical models for excitable systems may describe a wide range of biological phenomena. Among these phenomena, the most known and studied are certainly the two following ones: the generation and propagation of the nerve impulse along a nerve fiber and the generation and propagation of a cardiac pulse in cardiac cells. For both, following the seminal work [HH52], very detailed models known as conductance based models have been developed, describing the physio-biological mechanism leading to the generation and propagation of an action potential. These physiological models are quite difficult to handle mathematically and phenomenological models have been proposed. These models describe qualitatively the generation and propagation of an action potential in excitable systems. For instance, the Morris-Lecar model for the nerve impulse and the Fitzhugh-Nagumo model for the cardiac potential. In the present chapter, we will consider two phenomenological models: the Barkley and the Mitchell-Schaeffer models. Mathematically, they consist in a degenerate system of Partial Differential Equations (PDEs) driven by a stochastic term, often referred to as noise. More precisely, they model may be written

$$\left\{\begin{array}{lcl} \mathrm{d}u & = & [\nu\Delta u + \frac{1}{\varepsilon}f(u,v)]\mathrm{d}t + \sigma \mathrm{d}W, \\ \mathrm{d}v & = & g(u,v)\mathrm{d}t, \end{array}\right. \tag{6.1}$$

on $[0,T] \times D$, where D is a regular bounded open set of $\mathbb{R}^2$ or $\mathbb{R}^3$. This system is completed with boundary and initial conditions. W is a colored Gaussian noise source which will be defined more precisely later. System (6.1) is degenerate in two ways: there is no spatial operator such as the Laplacian neither noise source in the equation on v. All the considered models have the features of classical stochastic PDEs for excitable systems. The general structure of f and g is also typical of excitable dynamics. In particular, in the models that we will consider, the neutral curve $f(u,v) = 0$ when v is held fixed is cubic in shape.

To achieve our first aim, that is to numerically compute a solution of system (6.1), we work with a numerical scheme based on finite difference discretization in time and finite element method in space. The choice of finite element discretization in space has been directed by two considerations. The first is that this method fits naturally to a general spatial domain: we want to investigate the behavior of solutions to (6.1) on domains with various geometry. The second is that it allows to implement numerically the scheme using popular software such as the finite element software FreeFem++ or equivalent. The discretization of SPDE by finite differences in time and finite elements in space has been considered by several authors in theoretical studies, see for example [DP09, CYY07, LT12, Wal05]. Other methods of discretization are considered for example in [ANZ98, GMV12, Jen09, JR10, KLL10, LT10, Yan05]. These methods are based on finite difference discretization in time coupled either to finite difference in space or to the Galerkin spectral method, or to the finite element method on the integral formulation of the evolution equation. We emphasize that we do not consider in this chapter a Galerkin spectral method or exponential integrator, that is, roughly speaking, we neither use the spectral decomposition of the solution of (6.1) according to a Hilbert basis of $L^2(D)$ (or an other Hilbert space related to D) nor the semigroup attached to the linear operator (the Laplacian in (6.1)), in order to build our scheme. We only use the variational version of the finite element method in order to fit to commonly used finite elements method for deterministic PDEs. Moreover, the present chapter is more numerically oriented than the above cited papers, in the spirit of [Sha05]. In [Sha05], the author numerically analyzes the effect of noise on excitable systems thanks to a Galerkin spectral method of discretization on the square. In the present chapter, we pursue the same objective using the finite element method instead of the Galerkin spectral one. We believe that the finite element method is easier to adapt to various spatial domains. Let us notice that a discretization scheme for SPDEs driven by white noise for spatial domains of dimension greater or equal to 2 may lead to non trivial phenomena, see [HRW12]. Considering colored noises may also be seen as a way to circumvent these difficulties.

As is well known, one can consider two types of errors related to a numerical scheme for stochastic evolution equations: the strong error and the weak error. The strong error for discretization we consider has been analyzed for one dimensional spatial domains (line segments) in [Wal05]. The weak error for more general spatial domains, of dimension 2 or 3 for example, has been considered in [DP09]. In the present chapter, we prefer to consider the strong error of convergence of our scheme because we want to investigate numerally pathwise properties of the model. Working with spatial domains of dimension 2 or 3, we show that the strong order of convergence of the considered method for a class of linear stochastic equations is twice less than the weak order obtained in [DP09]. This is what is expected since this same duality between weak and strong order holds for the discretization of finite dimensional stochastic differential equations (SDEs). Thanks to the spatial regularity of the considered noise, the proof we provide follows classical arguments used to analyze the error introduced by the deterministic finite element method.

Our motivation for considering systems such as (6.1) comes from biological considerations. In the cardiac muscle, tachyarrhythmia are disturbances of the heart rhythm in which the heart beating rate is abnormally increased. This is a major trouble of the cardiac rhythm since it may lead to rapid loss of consciousness and to death. As explained in [Hin02, JC06], the vast majority of tachyarrhythmia are perpetuated by a reentrant mechanism. It is well known that deterministic excitable systems of type (6.1) are able to generate sustained reentrant patterns such as spiral or meander, see for example [Kee80, BKT90]. We show numerically that reentrant patterns may be generated and perpetuated only by the presence of noise. We perform the simulations on the Barkley model whose deterministic version has been intensively studied in [BKT90, Bar92, Bar94] and the model of Mitchell-Schaeffer which allows to get more realistic shape for the action potential in cardiac cells [MS03, BCF+10]. For the Barkley model, similar experiments are presented in [Sha05] where Galerkin spectral method is used as simulation scheme on a square domain. In our simulations, done on a square with periodic conditions or on a smoothed cardioid, we observe two kinds of reentrant patterns due to noise: the first may be seen as a scroll wave phenomenon whereas the second corresponds to spiral phenomenon. Both phenomena may be regarded as sources of tachyarrhythmia since in both cases, areas of the spatial domain are successively activated by the same wave which re-enters in the region.

All the simulations in the present chapter have been performed using the FreeFem++ finite element software, see [HLHOP]. This software offers the advantage to provide the mesh of the domain, the corresponding finite element basis and to solve linear problems related to the finite element discretization of the model on its own. The originality of the present work is to use this software to simulate stochastic PDEs.

Let us emphasize that the generic model (6.1) is endowed with a timescale parameter ε. The presence of this parameter is fundamental for the observation of traveling waves in the system: ε enforces the system to be either quiescent or excited with a sharp transition between the two states. Moreover, the values of the timescale parameter ε and the strength of the noise σ appear to be of first importance to obtain reentrant patterns. This fact is also pointed out by our numerical bifurcation analysis. Let us mention that noise induced phenomena have been studied in [BG06] for finite dimensional systems of stochastic differential equations. The theoretical study of slow-fast SPDEs, through averaging methods, has been considered in [Bre12, CF09, WR12] for SPDEs.

In forthcoming work, we plan to address the effect of noise on deterministic periodic forcing of the Barkley and Mitchell-Schaeffer models. We expect to observe as in [TJ10] for the one dimensional case, the annihilation by weak noise of the propagation of some waves initiated by deterministic periodic forcing. We also want to investigate stochastic resonance phenomena in such a situation. On a theoretical point of view, we intend to derive the strong order of convergence of the discretization method used in the present chapter for non-linear equations and systems of equations such as the FitzHugh-Nagumo, Barkley or Mitchell-Schaeffer models but also on simplified conductance based models.

The remainder of the chapter is organized as follows. In Section 6.2, we begin by the precise definition of the noise source in system (6.1) and present its finite element discretization. In Section 6.3, we introduce a discretization scheme based on finite element in space for a stochastic heat equation and we estimate the strong order of convergence. Then we apply the method to the Fitzhugh-Nagumo model. In Section 6.4, we investigate the influence of noise on the Barkley and Mitchell-

Schaeffer models. We show that noise may initiate reentrant patterns which are not observable in the deterministic case. We also provide numerical bifurcation diagrams between the noise intensity σ and the time-scale ε of the models. At last, some proofs are postponed to the Appendix.

6.2 Finite element discretization of Q-Wiener processes.

6.2.1 Basic facts on Q-Wiener processes

Let D be an open bounded domain of $\mathbb{R}^d$, $d = 2$ or 3, containing the origin and with polyhedral frontier. We denote by $L^2(D)$ the set of square integrable measurable functions with respect to the Lebesgue measure on $\mathbb{R}^d$. Writing H for $L^2(D)$, we recall that H is a real separable Hilbert space and we denote its usual scalar product by $(\cdot,\cdot)$ and the associated norm by $\|\cdot\|$. They are respectively given by

$$\forall(\phi_1,\phi_2) \in H \times H, \quad (\phi_1,\phi_2) = \int_D \phi_1(x)\phi_2(x)\mathrm{d}x, \quad \|\phi_1\| = \left(\int_D \phi_1(x)^2\mathrm{d}x\right)^{\frac{1}{2}}.$$

Let Q be a non-negative symmetric operator on H. Let us recall the definition of a Q-Wiener process on H which can be found in [PZ07], Section 4.4, as well as the basic properties of such a process.

Definition 6.2.1. *There exists a probability space $(\Omega,\mathcal{F},\mathbb{P})$ on which we can define a stochastic process $(W_t^Q, t \in \mathbb{R}_+)$ on H such that*

- *For each $t \in \mathbb{R}_+$, W_t^Q is a H-valued random variable.*
- *W^Q starts at 0 at time 0: $W_0^Q = 0_H$, $\mathbb{P}$-a.s.*
- *$(W_t^Q, t \in \mathbb{R}_+)$ is a Lévy process, that is, it is a process with independent and stationary increments:*
    - *Independent increments: for a sequence $t_1,\dots,t_n$ of strictly increasing times, the random variables $W_{t_2}^Q - W_{t_1}^Q, \dots, W_{t_n}^Q - W_{t_{n-1}}^Q$ are independent.*
    - *Stationary increments: for two times $s < t$, the random variable $W_t^Q - W_s^Q$ has same law as W_{t-s}^Q.*
- *$(W_t^Q, t \in \mathbb{R}_+)$ is a Gaussian process: for any $t \in \mathbb{R}_+$ and any $\phi \in H$, (W_t^Q,ϕ) is a real centered Gaussian random variable with variance $t(Q\phi,\phi)$.*
- *$(W_t^Q, t \in \mathbb{R}_+)$ is pathwise continuous $\mathbb{P}$-almost surely.*

We recall the definition of non-negative symmetric linear operator on H admitting a kernel.

Definition 6.2.2. *A non-negative symmetric linear operator $Q : H \to H$ is a linear operator defined on H such that*

$$\forall(\phi_1, \phi_2) \in H \times H, \quad (Q\phi_1, \phi_2) = (Q\phi_2, \phi_1), \quad (Q\phi_1, \phi_1) \geq 0.$$

Let q be a real valued integrable function on $D \times D$ such that

$$\forall(x, y) \in \overline{D} \times \overline{D}, \quad q(x, y) = q(y, x),$$
$$\forall M \in \mathbb{N}, \forall x_i, y_j \in \overline{D}, \forall a_i \in \mathbb{R}, i, j = 1, \dots M, \quad \sum_{i,j=1}^{M} q(x_i, y_j)a_i a_j \geq 0.$$

that is q is symmetric and non-negative definite on $\overline{D} \times \overline{D}$. We say that Q has the kernel q if

$$\forall \phi \in H, \forall x \in D, \quad Q\phi(x) = \int_D \phi(y)q(x, y)\mathrm{d}y.$$

Let $Q : H \to H$ be a non-negative symmetric operator with kernel q. Then Q is a trace class operator whose trace is given by

$$\mathrm{Tr}(Q) = \int_D q(x, x)\mathrm{d}x.$$

For examples of kernels and basic properties of symmetric non-negative linear operators on Hilbert spaces, we refer to [PZ07], Section 4.9.2 and Appendix A. Let us now state clearly our assumptions on the operator Q.

Assumption 6.2.1. *The operator Q is a non-negative symmetric operator with kernel q given by*

$$\forall(x, y) \in \overline{D} \times \overline{D}, \quad q(x, y) = C(x - y),$$

where C belongs to $\mathcal{C}^3(\overline{D})$ and is an even function on $\overline{D}$ satisfying:

$$\forall M \in \mathbb{N}, \forall x_i, y_j \in \overline{D}, \forall a_i \in \mathbb{R}, i, j = 1, \dots M, \quad \sum_{i,j=1}^{M} C(x_i - y_j)a_i a_j \geq 0$$

Particularly, $\nabla C(0) = 0$ and $x \mapsto \frac{C(x)-C(0)}{|x|^2}$ is bounded on a neighborhood of zero.

For $x \in \overline{D}$ and $t \in \mathbb{R}_+$, one can show, see [PZ07], Section 4.4, that we can define W_t^Q at the point x such that the process $(W_t^Q(x), (t,x) \in \mathbb{R}_+ \times \overline{D})$ is a centered Gaussian process with covariance between the points (t,x) and (s,y) given by

$$\mathbb{E}(W_t^Q(x) W_s^Q(y)) = t \wedge s \ q(x,y).$$

In this case, the correlations in time are said to be white whereas the correlations in space are colored by the kernel q.

Proposition 6.2.1. *Under Assumption 6.2.1, the process $(W_t^Q(x), (t,x) \in \mathbb{R}_+ \times \overline{D})$ has a version with continuous paths in space and time.*

Proof. This is an easy application of the Kolmogorov-Chentsov test, see [DPZ92] Chapter 3, Section 3.2. Note that the regularity of C is important to get the result. $\square$

Remark 6.2.1. *The more the kernel q smooth is, the more the Wiener process regular is. For example, let $\vec{\iota} \in \mathbb{R}^d$ and define $f(x) = \vec{\iota} \cdot \mathrm{Hess}C(x)\vec{\iota}$ for $x \in \overline{D}$. Suppose that f is a twice differentiable function. Then, one can show that there exists a probability space on which $(W_t^Q, t \in \mathbb{R}_+)$ is twice differentiable in the direction $\vec{\iota}$.*

Remark 6.2.2. *Let us assume that there exists a constant α and a (small) positive real δ such that*

$$\forall y \in \overline{D}, \quad |C(0) - C(y)| \leq \alpha |y|^{2+\delta}.$$

Using the Kolmogorov-Chentsov continuity theorem, one can show that the process

$$(W_t^Q(x), (t,x) \in \mathbb{R}_+ \times \overline{D})$$

has a modification which is γ_1-Hölder in time for all $\gamma_1 \in \left(0, \frac{1}{4}\right)$ and γ_2-Hölder in space for all $\gamma_2 \in \left(0, 1 + \frac{\delta}{2}\right)$. Thus if $\delta > 0$, by Rademacher theorem, for $\gamma_2 = 1$, this version is almost everywhere differentiable on $\overline{D}$. This is another way to obtain regularity in space for W^Q without using another probability space.

Remark 6.2.3. *For our purpose, on may also replace the assumption $C \in \mathcal{C}^3(\overline{D})$ by the weaker condition $C \in \mathcal{C}^{2+\varepsilon}(\overline{D})$ for some (small) positive ε.*

6.2.2 Finite element discretization

In this part, we assume that Q satisfies Assumption 6.2.1. Let us present our approximation of the Q-Wiener process W^Q. We begin with the discretization of

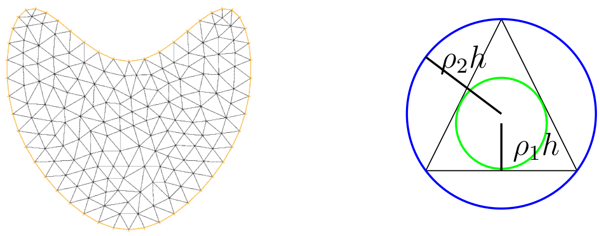


Figure 6.1: Meshing (triangulation) of a domain and illustration of (6.2)

the domain D. Let $\mathcal{T}_h$ be a family of triangulations of the domain D by triangles ($d = 2$) or tetrahedra ($d = 3$). The size of $\mathcal{T}_h$ is given by

$$h = \max_{T \in \mathcal{T}_h} h(T),$$

where $h(T) = \max_{x,y \in T} |x - y|$ is the diameter of the element T. We assume that there exist two positive constants ρ_1 and ρ_2 such that

$$\forall T \in \mathcal{T}_h, \quad \exists x \in T, \quad T \subset \mathrm{B}(x, \rho_2 h), \tag{6.2}$$

where $\mathrm{B}(x, r)$ stands for the euclidean ball centered at x with radius r. We assume further that this triangulation is admissible as in Figure 6.1 where a triangulation is displayed and the property (6.2) is illustrated. In the present work, we consider two kinds of finite elements: the Lagrangian P0 and P1 finite elements. However, the method could be adapted to other finite elements. The basis associated to the P0 finite element method is

$$\mathcal{B}_{0,\mathcal{T}_h} = \{1_T, T \in \mathcal{T}_h\},$$

where the function 1_T denotes the indicator function of the element T. Let $\{P_i, 1 \leq i \leq N_h\}$ be the set of all the nodes associated to the triangulation $\mathcal{T}_h$. The basis for the P1 finite element method is given by

$$\mathcal{B}_{1,\mathcal{T}_h} = \{\psi_i, 1 \leq i \leq N_h\},$$

where ψ_i is the continuous piecewise affine function on D defined by $\psi_i(P_j) = \delta_{ij}$ (Kronecker symbol) for all $1 \leq i, j \leq N_h$.

Definition 6.2.3. *The P0 approximation of the noise W^Q is given for $t \in \mathbb{R}_+$ by*

$$W_t^{Q,h,0} = \sum_{T \in \mathcal{T}_h} W_t^Q(g_T) 1_T, \tag{6.3}$$

where g_T is the center of gravity of T. The P1 approximation is

$$W_t^{Q,h,1} = \sum_{i=1}^{N_h} W_t^Q(P_i)\psi_i. \tag{6.4}$$

We will also consider the following alternative choice for the P0 discretization

$$W_t^{Q,h,0_a} = \sum_{T\in\mathcal{T}_h} \frac{1}{|T|}(W_t^Q, 1_T)1_T. \tag{6.5}$$

$W^{Q,h,0_a}$ corresponds to an orthonormal projection on P0. These approximations are again Wiener processes as stated in the following proposition.

Proposition 6.2.2. *For $i \in \{0, 0_a, 1\}$ the stochastic processes $(W_t^{Q,h,i}, t \in \mathbb{R}_+)$ are centered $Q^{h,i}$-Wiener processes where, for $\phi \in H$*

$$\begin{aligned} Q^{h,0}\phi &= \sum_{T,S\in\mathcal{T}_h} (1_T, \phi) q(g_T, g_S) 1_S, \\ Q^{h,0_a}\phi &= \sum_{T,S\in\mathcal{T}_h} (1_T, \phi) \frac{(Q1_T, 1_S)}{|T||S|} 1_S \end{aligned}$$

and

$$Q^{h,1}\phi = \sum_{i,j=1}^{N_h} (\psi_i, \phi) q(P_i, P_j)\psi_j.$$

Proof. The fact that for $i \in \{0, 0_a, 1\}$ the stochastic processes $(W_t^{Q,h,i}, t \in \mathbb{R}_+)$ are Wiener processes is a direct consequence of their definition as linear functionals of the Wiener process $(W_t^Q, t \in \mathbb{R}_+)$, see Definition 6.2.3. The corresponding covariance operators are obtained by computing the quantity

$$\mathbb{E}((W_1^{Q,h,i}, \phi_1)(W_1^{Q,h,i}, \phi_2))$$

for $i \in \{0, 0_a, 1\}$ and $\phi_1, \phi_2 \in H$ (the details are left to the reader). $\square$

The P0 approximation (6.5) of the noise has been considered for white noise in dimension 2 in [CYY07]. White noise corresponds to $Q = \mathrm{Id}_H$. In the white noise case, the associated Wiener process is not at all regular in space (the trace of Q is infinite in this case). In the present paper, we work with trace class operators and thus with noises which are regular in space. Notice that discretization schemes for SPDE driven by white noise for spatial domains of dimension greater or equal to 2 may lead to non trivial phenomena. In particular, usual schemes may not converge to the desired SPDE, see [HRW12].

Theorem 6.2.1 (A global error). *For any $\tau \in \mathbb{R}_+$ and $i \in \{0, 0_a, 1\}$ we have*

$$\mathbb{E}\left(\sup_{t\in[0,\tau]} \|W_t^Q - W_t^{Q,h,i}\|^2\right) \leq K\tau h^2$$

for a deterministic constant K depending only on $|D|$.

Proof. The proof is postponed to Appendix 6.A. □

Let us comment the above result. Let us take, as it will be the case in the numerical experiments, the following special form for the kernel

$$\forall x \in D, \quad C_\xi(x) = \frac{a}{\xi^2} e^{-\frac{b}{\xi^2}|x|^2}$$

for three positive real numbers a, b, ξ. This is a so-called Gaussian kernel. To a particular ξ-dependent kernel C_ξ, we associate the corresponding ξ-dependent covariance operator Q_ξ. Then, if we adapt the proof of Theorem 6.2.1 in this particular case, we see that

$$\mathbb{E}\left(\sup_{t\in[0,\tau]} \|W_t^{Q_\xi} - W_t^{Q_\xi,h,i}\|^2\right) = \mathrm{O}\left(\tau\frac{h^2}{\xi^4}\right)$$

for any $\tau \in \mathbb{R}_+$. Thus, when ξ goes to zero, this estimation becomes useless since the right hand-side goes to infinity. In fact, when ξ goes to zero, C_ξ converges in the distributional sense to a Dirac mass, and W^{Q_ξ} tends to a white noise which is, as mentioned before, an irregular process. In particular, the white noise does not belong to H and this is why our estimation is no longer useful in this case. Let us mention that for white noise acting on steady PDEs and on particular domains (square and disc), the error considered in Theorem 6.2.1 have been studied in [CYY07]: the regularity of the colored noised improved these estimates in our case. We also remark that the proof of Theorem 6.2.1 does not rely on the regularity of the functions of the finite element basis of P0 or P1 here. The key points are that $C \in \mathcal{C}^2(D)$ is even and that $\sum_i \phi_i = 1$, where $\{\phi_i\}$ corresponds to the finite element basis.
To conclude this section, we display some simulations. In Figure 6.2 are simulations of the noise $W_1^{Q_\xi}$ with covariance kernel defined by

$$\forall (x,y) \in D \times D, \quad q_\xi(x,y) = C_\xi(x-y) = \frac{1}{4\xi^2} e^{-\frac{\pi}{4\xi^2}|x-y|^2}, \tag{6.6}$$

where $\xi > 0$. We use the same kernel as in [Sha05] for comparison purposes. As already mentioned, when ξ goes to zero, the considered colored noise tends to

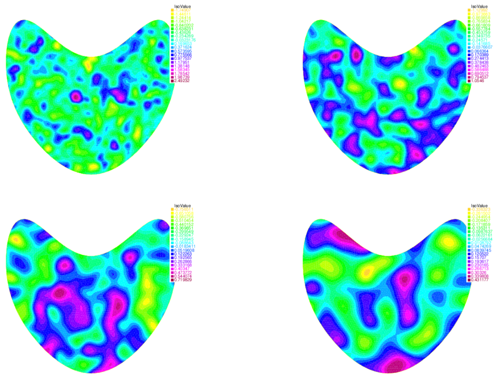


Figure 6.2: Simulations of $W_1^{Q_\xi}$ with $\xi = 1, 1.5, 2, 3$ with P1 finite elements.

a white noise. On the contrary, when ξ increases, the correlation between two distinct areas increases as well. This property is illustrated in Figure 6.2. In other words, ξ is a parameter which allows to control the spatial correlation.
In these simulations, we have discretized $W^{Q_\xi,h,1}$ with the P1 discretization which reads

$$W_1^{Q_\xi,h,1} = \sum_{i=1}^{N_h} W_1^{Q_\xi}(P_i)\psi_i.$$

We remark that the family $\{W_1^{Q_\xi}(P_i), 1 \leq i \leq N_h\}$ is a centered Gaussian vector with covariance matrix $(q_\xi(P_i,P_j))_{1\leq i,j\leq N_h}$. Using some basic linear algebra, it is not difficult to simulate a realization of this vector and to project it on the P1 finite element basis to obtain Figure 6.2.
We now propose a log-log graph to illustrate the estimate of Theorem 6.2.1. In the case where $\overline{D}$ is the square $[0,1]\times[0,1]$, let us consider the kernel q given by:

$$\forall (x,y) \in \overline{D}\times\overline{D}, \quad q(x,y) = f_{k_0p_0}(x)f_{k_0p_0}(y)$$

where for two given integers $k_0, p_0 \geq 1$, $f_{k_0p_0}(x) = 2\sin(k_0\pi x_1)\sin(p_0\pi x_2)$ (if $x = (x_1,x_2)\in\overline{D}$). Then, the covariance operator is given by $Q\phi = (\phi, f_{k_0p_0})f_{k_0p_0}$ for any $\phi \in L^2(\overline{D})$. Moreover, one can show that

$$\forall t \geq 0, \quad W_t^Q = \beta_t f_{k_0p_0}$$

for some real-valued Brownian motion β. All the calculations are straightforward in this setting. Suppose, in the finite element setting, that the square $\overline{D}$ is covered by $2N^2$ triangles, $N\in\mathbb{N}$. For any $N\in\mathbb{N}$, we denote by $W_1^{Q,N,0}$ the P0 approximation of W_1^Q given by (6.3). We show in Figure 6.3 the log-log graph of the (discrete)

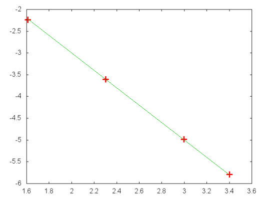


Figure 6.3: Log-log graph of $N \mapsto \mathrm{err}_N$ for $N = 5, 10, 20, 30$. In green is the comparison with a line of slope -2.

function:

$$N \mapsto \mathrm{err}_N = \mathbb{E}(\|W_1^Q - W_1^{Q,N,0}\|^2).$$

According to Theorem 6.2.1, we should have $\mathrm{err_N} = \mathrm{O}\left(\frac{1}{2N^2}\right)$. This result is recovered numerically in Figure 6.3.

6.3 Space-time numerical scheme

In this section, we first present our numerical scheme. The considered space-time discretization is based on the Euler-Maruyama scheme in time and on finite elements in space. In this section and in the next section, we will use the following notations. Let us fix a time horizon T. For $N \geq 1$ we define a time step $\Delta t = \frac{T}{N}$ and denote by $(u_n, v_n)_{0\leq n\leq N}$ a sequence of approximations of the solution of (6.1) at times $t_n = n\Delta t$, $0 \leq n \leq N$. The scheme, semi-discretized in time, is based on the following variational formulation, for $n \in \{0, \ldots, N-1\}$:

$$\left\{ \begin{array}{rcl} \left(\frac{u_{n+1}-u_n}{\Delta t}, \psi\right) + \kappa(\nabla u_{n+1}, \nabla\psi) & = & \frac{1}{\varepsilon}(f_n, \psi) + \frac{\sigma}{\sqrt{\Delta t}}(W_{n+1}^Q - W_n^Q, \psi), \\ \left(\frac{v_{n+1}-v_n}{\Delta t}, \psi\right) & = & (g_n, \psi) \end{array} \right. \tag{6.7}$$

for ψ in an appropriate space of test functions. Here, f_n and g_n correspond to approximations of the reaction terms f and g in (6.1). The way we compute f_n and g_n is detailed in the sequel for each considered model. W_n^Q is an appropriate approximation of $W_{t_n}^Q$ based on one of the discretization proposed in Definition 6.2.3.

In Subsection 6.3.1, we consider the strong error in the case of a linear stochastic partial differential equation driven by a colored noise to study the accuracy of the finite element discretization. We obtain that the strong order of convergence of the scheme is twice less than the weak order obtained in [DP09], as expected. In Subsection 6.3.2, we implement the scheme for the Fitzhugh-Nagumo model with a colored noise since this model is one of the most used phenomenological models in cardiac electro-physiology, see the seminal work [Fit69] and the review [LGONSG04].

6.3.1 Linear parabolic equation with additive colored noise

Let us consider the following linear parabolic stochastic equation on $(0,T)\times D$

$$\left\{\begin{array}{rcl} \mathrm{d}u_t &=& Au_t\mathrm{d}t+\sigma\mathrm{d}W_t^Q, \\ u_0 &=& \zeta. \end{array}\right. \tag{6.8}$$

Remember that $H=L^2(D)$ is a separable Hilbert space with the scalar product and the corresponding norm respectively denoted by $(\cdot,\cdot)$ and $\|\cdot\|$. We assume that W^Q is a Q-Wiener process with an operator Q which satisfies Assumption 6.2.1. Let ζ be a $H^2(D)$-valued random variable. We impose the following condition on the operator A in (6.8).

Assumption 6.3.1. *The operator $-A$ is a positive self-adjoint linear operator on H whose domain is dense and compactly embedded in H.*

It is well known that the spectrum of $-A$ is made up of an increasing sequence of positive eigenvalues $(\lambda_i)_{i\geq 1}$. The corresponding eigenvectors $\{w_i, i\geq 1\}$ form a Hilbert basis of H. The following proposition states that problem (6.8) is well posed.

Proposition 6.3.1. *Equation (6.8) has a unique mild solution:*

$$u_t=e^{At}\zeta+\sigma\int_0^t e^{A(t-s)}\mathrm{d}W_s^Q,$$

Moreover u is continuous in time and $u_t\in H$ for all $t\in[0,T]$, $\mathbb{P}$-a.s.

Proof. This result is a direct consequence of Theorem 5.4 of [DPZ92], Assumptions 6.3.1 and 6.2.1. □

The domain of $(-A)^{\frac{1}{2}}$ is the set

$$\left\{u=\sum_{i\geq 1}(u,w_i)w_i,\quad \sum_{i\geq 1}\lambda_i(u,w_i)^2<\infty\right\},$$

that we denote here by V. It is continuously and densely embedded in H. The V-norm is given by $|u|=\sqrt{-(Au,u)}$ for all $u\in V$. We define a coercive continuous bilinear form a on $V\times V$ by

$$a(u,v)=-(Au,v).$$

For $h > 0$, let V_h be a finite dimensional subset of V with the property that for all $v \in V$, there exists a sequence of elements $v_h \in V_h$ such that $\lim_{h\to 0} \|v - v_h\| = 0$. For an element u of V, we introduce its orthogonal projection on V_h and denote it $\Pi_h u$. It is defined in a unique way by

$$\Pi_h u \in V_h \quad \text{and} \quad \forall v_h \in V_h, \quad a(\Pi_h u - u, v_h) = 0. \tag{6.9}$$

Let I_h be the dimension of V_h. Notice that there exists a basis $(w_{i,h})_{1\le i\le I_h}$ of V_h orthonormal in H with the following property: for each $1 \le i \le I_h$, there exists $\lambda_{i,h}$ such that

$$\forall v_h \in V_h, \quad a(v_h, w_{i,h}) = \lambda_{i,h}(v_h, w_{i,h}),$$

(see [RT83], Section 6.4). The family $(\lambda_{i,h})_{1\le i\le I_h}$ is an approximating sequence of the family of eigenvalues $(\lambda_i)_{i\ge 1}$ so that

$$\lambda_{i,h} \ge \lambda_i, \quad \forall 1 \le i \le I_h.$$

We study the following numerical scheme to approximate equation (6.8) defined recursively as follows. For u_0 given in V_h, find $(u_n^h)_{0\le n\le N}$ in V_h such that for all $n \le N - 1$

$$\left\{ \begin{array}{rcl} \frac{1}{\Delta t}(u_{n+1}^h - u_n^h, v_h) + a(u_{n+1}^h, v_h) & = & \frac{\sigma}{\Delta t}(W_{n+1}^{Q,h} - W_n^{Q,h}, v_h) \\ u_0^h & = & u_0 \end{array} \right. \tag{6.10}$$

for all $v_h \in V_h$ where $W_n^{Q,h}$ is an appropriate approximation of $W_{n\Delta t}^Q$ in V_h. The approximation error of the scheme can be written as the sum of two errors.

Definition 6.3.1. *The discrete error introduced by the scheme (6.10) is defined by $E_n^h = e_n^h + p_n^h$ where*

$$e_n^h = u_n^h - \Pi_h u_{t_n}, \quad p_n^h = \Pi_h u_{t_n} - u_{t_n} \tag{6.11}$$

for $0 \le n \le N$. The consistency error ε_n^h is defined, for $n \in \{0, \dots, N\}$ and $v_h \in V_h$ by

$$\begin{aligned}(\varepsilon_n^h, v_h) = {} & \frac{\sigma}{\Delta t}(W_{n+1}^{Q,h} - W_n^{Q,h}, v_h) - \frac{\sigma}{\Delta t}(W_{(n+1)\Delta t}^Q - W_{n\Delta t}^Q, v_h) \\ & + \frac{1}{\Delta t}\int_{n\Delta t}^{(n+1)\Delta t} a(u_s, v_h)\mathrm{d}s - a(u_{(n+1)\Delta t}, v_h) + \frac{1}{\Delta t}((\mathrm{I} - \Pi_h)(u_{(n+1)\Delta t} - u_{n\Delta t}), v_h).\end{aligned}$$

In this definition, I is the identity operator on H. For $n \in \{0, \dots, N\}$, the error e_n^h is the difference between the approximated solution given by the scheme and the elliptic projection on V_h of the exact solution at time $n\Delta t$. The error p_n^h is the difference between the exact solution and its projection on V_h at time $n\Delta t$. The following result gives the error e_n^h with respect to the previous consistency errors $(\varepsilon_k^h)_{0\le k\le n-1}$.

Theorem 6.3.1. *For $h > 0$ and $n \in \{0, \dots, N-1\}$ we have*

$$e_n^h = \sum_{i=1}^{I_h} e^{-\mu_{i,h,\Delta t} n\Delta t}(e_0^h, w_{i,h})w_{i,h} + \Delta t \sum_{i=1}^{I_h}\sum_{k=0}^{n-1} e^{-\mu_{i,h,\Delta t}(n-k)\Delta t}(\varepsilon_k^h, w_{i,h})w_{i,h} \quad (6.12)$$

with, for $1 \le i \le I_h$ and $h, \Delta t > 0$, $\mu_{i,h,\Delta t} = \frac{\log(1+\Delta t\lambda_{i,h})}{\Delta t}$.

For all $1 \le i \le I_h$ and $h, \Delta t > 0$, we have $\frac{1}{1+\Delta t\lambda_{i,h}} < 1$ such that the scheme (6.10) is stable.

Proof. The proof is postponed to Appendix 6.B. □

To bound the consistency error, we have to choose our approximation in H specifically. This imposes to choose the space V explicitly. We set $V = H_0^1(D)$ and V_h a space of P1 finite elements, see Section 6.2.2. In this P1 case, for $n \in \{0, \dots, N\}$, we set $W_n^{Q,h} = W_{n\Delta t}^{Q,h,1}$ defined by Definition 6.2.3.
For technical reasons, we make the following assumption.

Assumption 6.3.2. *We assume that*

$$\sum_{p=0}^{2}\int_0^T \mathrm{Tr}\left(\nabla^p e^{As}Q^{\frac12}\left(\nabla^p e^{As}Q^{\frac12}\right)^*\right)\mathrm{d}s < \infty.$$

Remark 6.3.1. *The above assumption implies that $\sup_{t\in[0,T]}\mathbb{E}(\|u_t\|^2_{H^2(D)}) < \infty$ (see Theorem 5.20 of [DPZ92]). With $A = \Delta$ and $\mathcal{D}(A) = H^2(D) \cap H_0^1(D)$, Assumption 6.3.2 is fulfilled as soon as $\mathrm{Tr}((-\Delta)^{1+\delta}Q) < \infty$ for some $\delta > 0$. The latter estimate is satisfied since $C \in \mathcal{C}^3(\overline{D})$ according to Assumption 6.2.1.*

Theorem 6.3.2. *Let us assume that Assumptions 6.2.1, 6.3.1 and 6.3.2 are satisfied. Moreover, assume that we are in the P1 case: for $n \in \{0, \dots, N\}$, $W_n^{Q,h} = W_{n\Delta t}^{Q,h,1}$ defined by Definition 6.2.3. Then, there exists $\Delta t_0 > 0$ such that for all $n \in \{1, \dots, N\}$ and $\Delta t \in [0, \Delta t_0]$*

$$\mathbb{E}(\|e_n^h\|^2) \le \mathbb{E}(\|e_0^h\|^2) + K(\Delta t + h^2)$$

for a constant K depending only on $|D|$ and T.

Proposition 6.3.2. *The projection error p_n^h satisfies, for all $n \in \{1, \dots, N\}$:*

$$\sqrt{\mathbb{E}(\|p_n^h\|^2)} \le Kh \quad (6.13)$$

for a constant K depending only on $|D|$ and T.

Proof. Since we are working with P1 finite elements and the domain D is bounded and polyhedral, under the Assumption 6.3.2, the result is a direct consequence of Lemma 6.5.1 of [RT83]. □

Corollary 6.3.1. *Suppose* $\sqrt{\mathbb{E}(\|E_0^h\|^2)} = \mathrm{O}(h)$. *Assume that the hypotheses of Theorem 6.3.2 are satisfied. For all* $0 \leq n \leq N$ *and* $\Delta t \in [0, \Delta t_0]$

$$\sqrt{\mathbb{E}(\|E_n^h\|^2)} \leq K(h + \sqrt{\Delta t}), \tag{6.14}$$

where K *is a constant depending only on* T *and* $|D|$.

Before starting the proof of Theorem 6.3.2, we recall the weak order of convergence of the considered scheme obtained in [DP09] but under weaker assumptions. Since C is a twice differentiable even function on D, ΔC is a bounded function on D and therefore, according to [DP09] Theorem 3.1, for any bounded real valued twice differentiable function ϕ on $L^2(D)$, there exists a constant K depending only on T such that

$$|\mathbb{E}(\phi(u_N^h)) - \mathbb{E}(\phi(u_T))| \leq K(h^{2\gamma} + \theta^\gamma) \tag{6.15}$$

for a given $\gamma < 1$. In our situation, it is more natural to consider the strong error since we study pathwise behavior. For the method that we consider, estimates for the strong error have been obtained for one dimensional spatial domains and white noise in [Wal05]. Many papers exist for finite dimensional systems. Our estimate lies in between these two types of studies. We have a colored noise but our spatial domain may be of any dimension. We notice that the order of weak convergence (6.15) is twice the order of strong convergence (6.14), as for finite dimensional stochastic differential equations.
We begin the proof of Theorem 6.3.2 by estimating the error induced by the noise. From (6.12), we compare the discrete sum of the increments of the noise to stochastic integrals to obtain the order of convergence in time. Then, we compare the stochastic integrals to obtain the order of convergence in space. The proofs of the two following lemmas are postponed to Appendix 6.C.

Lemma 6.3.1. *Remember that, for* $n \in \{0, \dots, N\}$, $W_n^{Q,h} = W_{n\Delta t}^{Q,h,1}$ *defined in*

Definition 6.2.3. Let us define, for $n \in \{1, \dots, N\}$

$$A_n = \sum_{i=1}^{I_h} \sum_{k=0}^{n-1} e^{-\mu_{i,h,\Delta t}(n\Delta t - k\Delta t)} (W^Q_{(k+1)\Delta t} - W^Q_{k\Delta t}, w_{i,h}) w_{i,h},$$

$$B_n = \sum_{i=1}^{I_h} \left(\int_0^{n\Delta t} e^{-\mu_{i,h,\Delta t}(n\Delta t - s)} \mathrm{d}W^Q_s, w_{i,h} \right) w_{i,h},$$

$$C_n = \sum_{i=1}^{I_h} \sum_{k=0}^{n-1} e^{-\mu_{i,h,\Delta t}(n\Delta t - k\Delta t)} (W^{Q,h}_{k+1} - W^{Q,h}_{k}, w_{i,h}) w_{i,h},$$

$$D_n = \sum_{i=1}^{I_h} \left(\int_0^{n\Delta t} e^{-\mu_{i,h,\Delta t}(n\Delta t - s)} \mathrm{d}W^{Q,h,1}_s, w_{i,h} \right) w_{i,h}.$$

There exists a constant K_f such that for all $n \in \{1, \dots, N\}$

$$\mathbb{E}\left(\|A_n - B_n\|^2\right) \le K_f \mathrm{Tr}(Q)\Delta t,$$
$$\mathbb{E}\left(\|C_n - D_n\|^2\right) \le K_f \mathrm{Tr}(Q^{h,1})\Delta t.$$

For any $\Delta t_0 > 0$, there exists a constant K depending only on $|D|$ such that for all $n \in \{1, \dots, N\}$ and all $\Delta t \in [0, \Delta t_0]$,

$$\mathbb{E}(\|B_n - D_n\|^2) \le K h^2.$$

Remark 6.3.2. *The constant K_f is given by*

$$\sup_{x \in \mathbb{R}_+} \frac{\log(1+x) + \frac{1}{2}x^2 - x}{x(x+2)\log(1+x)}.$$

We go on with dealing with the terms in (6.12) which involve the real solution u of problem (6.8).

Lemma 6.3.2. *Let us define, for $n \in \{1, \dots, N\}$*

$$E_n = \sum_{i=1}^{I_h} \sum_{k=0}^{n-1} e^{-\mu_{i,h,\Delta t}(n-k)\Delta t} ((\mathrm{I} - \Pi_h)(u_{(k+1)\Delta t} - u_{k\Delta t}), w_{i,h}) w_{i,h},$$

$$F_n = \sum_{i=1}^{I_h} \sum_{k=0}^{n-1} e^{-\mu_{i,h,\Delta t}(n-k)\Delta t} \int_{k\Delta t}^{(k+1)\Delta t} a(u_{(k+1)\Delta t} - u_s, w_{i,h}) \mathrm{d}s\ w_{i,h}.$$

Under Assumption 6.3.2, there exists a constant K depending only on $|D|$ and T such that for all $n \in \{1, \dots, N\}$

$$\mathbb{E}(\|E_n\|^2) \le K h^2,$$
$$\mathbb{E}(\|F_n\|^2) \le K \Delta t.$$

Proof of Theorem 6.3.2. This is a consequence of the fact that, from (6.12), for all $n \in \{0, \ldots, N\}$:

$$\begin{aligned} e_n^h &= \sum_{i=1}^{I_h} e^{-\mu_{i,h,\Delta t}\Delta tn}(e_0^h, w_{i,h})w_{i,h} - \sigma A_n + \sigma C_n - E_n + F_n \\ &= \sum_{i=1}^{I_h} e^{-\mu_{i,h,\Delta t}\Delta tn}(e_0^h, w_{i,h})w_{i,h} - \sigma(A_n - B_n) - \sigma(B_n - D_n) + \sigma(C_n - D_n) - E_n + F_n. \end{aligned}$$

Then Lemmas 6.3.1 and 6.3.2 yield the result. □

In the end of this subsection, we illustrate this error estimate in a simple situation. We consider the domain $D = (0, l) \times (0, l)$ for $l > 0$. We set $A = \Delta$ and $\mathcal{D}(A) = H^2(D) \cap H_0^1(D)$. That is we consider the equation:

$$\begin{cases} \mathrm{d}u_t &= \Delta u_t \mathrm{d}t + \sigma \mathrm{d}W_t^Q, \quad \text{in } D, \\ u_t &= 0, \quad \text{on } \partial D, \\ u_0 &= 0, \quad \text{in } D \end{cases} \tag{6.16}$$

for $t \in \mathbb{R}_+$. W^Q is a $L^2(D)$-valued Q-Wiener process defined by Definition 6.2.1 and Q satisfies Assumption 6.2.1. Recall that in this case, there exists a $\varepsilon > 0$ $\mathrm{Tr}((-\Delta)^{1+\delta}Q) < \infty$ since C belongs to $\mathcal{C}^3(\overline{D})$ such that Assumption 6.3.2 is fulfilled. The initial condition is zero, hence the solution of the corresponding deterministic equation, without noise, is simply zero for all time.

Proposition 6.3.3. *Equation (6.16) has a unique mild solution such that $u_t \in L^2(D)$ for all $t \in [0, T]$, $\mathbb{P}$-almost surely. Moreover, u has a version with time continuous paths and such that, for any time $T > 0$:*

$$\sup_{t \in [0,T]} \mathbb{E}(\|u_t\|_{H^2(D)}^2) < \infty.$$

Proof. This result is a direct consequence of Theorem 5.20 of [DPZ92] and the fact that $\mathrm{Tr}((-\Delta)^{1+\delta}Q) < \infty$. □

We denote by $(e^{\Delta t}, t \geq 0)$ the contraction semigroup associated to the operator Δ. The mild solution to equation (6.16) is defined as the following stochastic convolution

$$u_t = \sigma \int_0^t e^{\Delta(t-s)} \mathrm{d}W_s^Q$$

for $t \in \mathbb{R}_+$, $\mathbb{P}$-almost-surely. In order to compute the expectation of the squared norm of u in $L^2(D)$ analytically and also as precisely as possible numerically, we define the Hilbert basis $(e_{kp}, k, p \geq 1)$ of $L^2(D)$ which diagonalizes the operator Δ defined on $\mathcal{D}(A)$. For $k, p \geq 1$ and $(x, y) \in D$

$$e_{kp}(x, y) = \frac{2}{l} \sin\left(\frac{k\pi}{l}x\right) \sin\left(\frac{p\pi}{l}y\right).$$

A direct computation shows that $\Delta e_{kp} = -\lambda_{kp} e_{kp}$ where $\lambda_{kp} = \frac{\pi^2}{l^2}(k^2 + p^2)$. In the basis $(e_{kp}, k, p \geq 1)$ of $L^2(D)$, the semigroup $(e^{\Delta t}, t \geq 0)$ is given by

$$e^{\Delta t}\phi = \sum_{k,p \geq 1} e^{-\lambda_{kp} t} (\phi, e_{kp}) e_{kp}$$

for $t \in \mathbb{R}_+$ and $\phi \in L^2(D)$. Then for any $t \in \mathbb{R}_+$ (c.f. Proposition 2.2.2 of [?])

$$\mathbb{E}(\|u_t\|^2) = \sigma^2 \int_0^t \mathrm{Tr}\left(e^{2\Delta s} Q\right) \mathrm{d}s = \sigma^2 \sum_{k,p \geq 1} \frac{1 - e^{-2\lambda_{kp} t}}{2\lambda_{kp}} (Q e_{kp}, e_{kp}).$$

In the sequel, we write $\Gamma_t = \mathbb{E}(\|u_t\|^2)$. The above series expansion can then be implemented and we can compare this result with $\mathbb{E}(\|u_n^h\|^2)$ which is computed thanks to Monte-Carlo simulations. The Monte-Carlo simulation of $\mathbb{E}(\|u_n^h\|^2)$ consists in considering $(u_n^{h,p})_{1 \leq p \leq P}$, $P \in \mathbb{N}$ a sequence of independent realizations of the scheme (6.10) and define

$$\Gamma_{n\Delta t}^{(P)} = \frac{1}{P} \sum_{p=1}^{P} \|u_n^{h,p}\|^2, \tag{6.17}$$

the approximation of Γ at time $n\Delta t$, $n \in \{0, \ldots, N\}$. We denote also by $\Gamma^{(P)}$ the continuous piecewise linear version of Γ. Figure 6.4 displays numerical simulations of the processes $(\Gamma_t, t \in \mathbb{R}_+)$ and $(\Gamma_t^{(P)}, t \in \mathbb{R}_+)$. The simulations are done with $l = 80$. Moreover the domain is triangulated with 5000 triangles giving a space step of about $h = 0.64$ and a number of vertices's of about 2600. For this simulation, we choose $P = 40$ which is not big but $\Gamma^{(40)}$ matches quite well with its corresponding theoretical version Γ, as expected by the law of large numbers. We remark also that for the same spatial discretization of the domain D, there is no particular statistical improvement to choose the P1 finite element basis instead of the P0.

6.3.2 Space-time discretization of the Fitzhugh-Nagumo model

We write the scheme for the Fitzhugh-Nagumo which is a widely used model of excitable cells, see [Fit69, LGONSG04]. The stochastic Fitzhugh-Nagumo model,

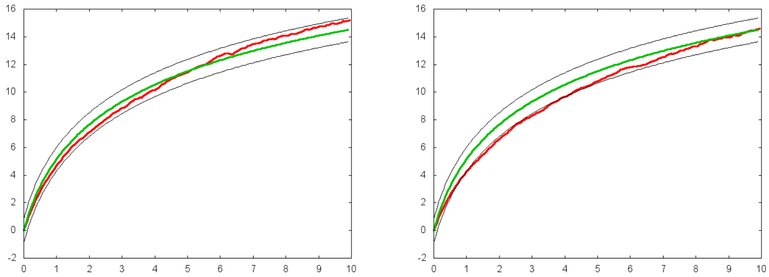


Figure 6.4: Simulations of $(\Gamma_t, t \in [0, 10])$ in green and its approximation $\Gamma^{(40)}$ in red computed with: the P0 (on the left) and P1 (on the right) approximations of the noise. For the simulation we choose the coefficient of correlation $\xi = 2$ with the kernel q_ξ defined by (6.6). The intensity of the noise is $\sigma = 0.15$. The time step is 0.05 whereas the space step is about 0.64. The two black curves are respectively Γ plus, respectively minus, the error introduced by the scheme which is expected to be of order $\sqrt{\Delta t} + h$ equals here to $\sqrt{0.05} + 0.64$.

abbreviated by FHN model in the sequel, consists in the following 2-dimensional system

$$\left\{ \begin{array}{rcl} \mathrm{d}u & = & [\kappa \Delta u + \frac{1}{\varepsilon} (u(1-u)(u-a) - v)]\mathrm{d}t + \sigma \mathrm{d}W^Q, \\ \mathrm{d}v & = & [u - v]\mathrm{d}t, \end{array} \right. \tag{6.18}$$

on $[0, T] \times D$. In the above system, $\kappa > 0$ is a *diffusion* coefficient, $\varepsilon > 0$ a *time-scale* coefficient, $\sigma > 0$ the intensity of the noise and $a \in (0, 1)$ a parameter. W^Q is a Q-Wiener process satisfying Assumption 6.2.1. System (6.18) must be endowed with initial and boundary conditions. We denote by u_0 and v_0 the initial conditions for u and v. Moreover we assume that u satisfies zero Neumann boundary conditions:

$$\forall t \in [0, T], \quad \frac{\partial u_t}{\partial \vec{n}} = 0, \quad \text{on } \partial D, \tag{6.19}$$

where ∂D denotes the boundary of D and $\vec{n}$ is the external unit normal to this boundary. Noisy FHN model and especially, FHN with white noise, have been extensively studied. We refer the reader to [BM08] where all the arguments needed to prove the following proposition are developed.

Proposition 6.3.4. *Let W^Q be a colored noise with Q satisfying Assumption 6.2.1. We assume that u_0 and v_0 are in $L^2(D)$, $\mathbb{P}$-almost surely. Then, for any time horizon T, the system (6.18) has a unique solution (u, v) defined on $[0, T]$ which is $\mathbb{P}$-almost surely in $\mathcal{C}([0, T], H) \times \mathcal{C}([0, T], H)$.*

The proof of this proposition relies on Itô Formula, see Chapter 1, Section 4.5 of [DPZ92], and the fact that the functional defined by

$$f(x) = x(1-x)(x-a), \quad \forall x \in \mathbb{R}$$

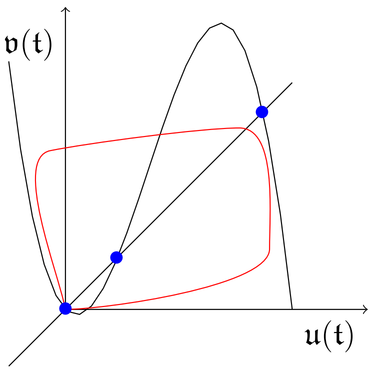


Figure 6.5: Phase portrait with nullclines of system (6.20) for $a = 0.1$ and $\varepsilon = 0.1$. The blue points correspond to the three equilibrium points of the system.

satisfies the inequality

$$(f(u) - f(v), u - v) \leq \frac{1 + a^2 - a}{3} \|u - v\|^2, \quad \forall (u, v) \in H \times H,$$

which implies that the map $f - \frac{1+a^2-a}{3}\mathrm{Id}$ is dissipative. The local kinetics of system (6.18), that is the dynamics in the absence of spatial derivative, is illustrated in Figure 6.5. It describes the dynamics of the system of ODEs

$$\left\{ \begin{array}{lcl} \mathrm{d}\mathfrak{u} & = & [\frac{1}{\varepsilon}\mathfrak{u}(1-\mathfrak{u})(\mathfrak{u}-a) - \mathfrak{v}]\mathrm{d}t, \\ \mathrm{d}\mathfrak{v} & = & [\mathfrak{u} - \mathfrak{v}]\mathrm{d}t, \end{array} \right. \tag{6.20}$$

when the initial condition $(\mathfrak{u}_0, \mathfrak{v}_0)$ is in $[0, 1] \times [0, 1]$.
We explicitly give the numerical scheme used to simulate system (6.18). Let us define the function k given by

$$k(x) = \frac{1}{\varepsilon}\left(-x^3 + x^2(1 + a)\right), \forall x \in \mathbb{R}.$$

This function is the non linear parts of the reaction term f. Other choices are possible like linearized f around 1 which is also a stable point or a which is unstable. We use the following semi-implicit Euler-Maruyama scheme

$$\left\{ \begin{array}{lcl} \frac{u_{n+1}-u_n}{\Delta t} & = & \kappa \Delta u_{n+1} - \frac{a}{\varepsilon} u_{n+1} + k(u_n) - v_{n+1} + \frac{\sigma}{\sqrt{\Delta t}} W^Q_{1,n+1}, \\ \frac{v_{n+1}-v_n}{\Delta t} & = & u_{n+1} - v_{n+1}, \end{array} \right. \tag{6.21}$$

where $(W^Q_{1,n})_{1 \leq n \leq N+1}$ is a sequence of independent Q-Wiener processes evaluated at time 1. Let $(H^*, \mathcal{B}(H^*), \tilde{\mathbb{P}})$ be chosen so that the canonical process has the same law as $W^Q_{1,n+1}$ under $\tilde{\mathbb{P}}$. Then, for a given $(u_n, v_n) \in H^1(D) \times H$, the equation

$$(\frac{1}{\Delta t} + \frac{a}{\varepsilon} + \frac{\Delta t}{1+\Delta t})u_{n+1} - \kappa \Delta u_{n+1} = k(u_n) - \frac{1}{1+\Delta t} v_n + \frac{\sigma}{\sqrt{\Delta t}} W^Q_{1,n+1}$$

has a unique weak solution u_{n+1} in $H^1(D)$, $\tilde{\mathbb{P}}$-almost surely. This fact follows from Lax-Milgram Theorem and a measurable selection theorem, see Section 5 of the survey [Wag80]. Therefore, without loss of generality, we may assume in this section that the probability space is $(\mathcal{C}([0,T], H^*), \mathcal{B}(\mathcal{C}([0,T], H^*)), \mathbb{P})$ such that under $\mathbb{P}$, the canonical process has the same law as W^Q. It is also possible to work with a completely implicit scheme with $k(u_{n+1})$ instead of $k(u_n)$ in (6.21).
We then consider the weak form for the first equation of (6.21). We get,

$$\left\{ \begin{array}{rcl} (\frac{1}{\Delta t} + \frac{a}{\varepsilon} + \frac{\Delta t}{1+\Delta t})(u_{n+1}, \psi) + \kappa(\nabla u_{n+1}, \nabla \psi) & = & (k(u_n), \psi) - \frac{1}{1+\Delta t}(v_n, \psi) + \frac{\sigma}{\sqrt{\Delta t}}(W^Q_{1,n+1}, \psi), \\ v_{n+1} - \frac{\Delta t}{1+\Delta t} u_{n+1} & = & \frac{1}{1+\Delta t} v_n \end{array} \right. \tag{6.22}$$

for all $\psi \in H^1(D)$. Let $h > 0$ and $(\psi_i, 1 \le i \le N_h)$ be the P1 finite element basis defined in Section 6.2. For $n \ge 0$, we define the vectors

$$\boldsymbol{u}_n = (u_{n,i})_{1 \le i \le N_h}, \quad \boldsymbol{v}_n = (v_{n,i})_{1 \le i \le N_h}, \quad \boldsymbol{W}^Q_{n+1} = (W^Q_{1,n+1}(P_i))_{1 \le i \le N_h},$$

which are respectively the coordinates of u_n, v_n and $W^Q_{1,n+1}$ w.r.t. the basis $(\psi_i, 1 \le i \le N_h)$. We also define the stiffness matrix $A \in \mathcal{M}_{N_h}(\mathbb{R})$ and the mass matrix $M \in \mathcal{M}_{N_h}(\mathbb{R})$ by

$$A_{ij} = (\nabla \psi_i, \nabla \psi_j), \quad M_{ij} = (\psi_i, \psi_j).$$

System (6.22) can be rewritten as

$$\begin{aligned} \begin{pmatrix} (\frac{1}{\Delta t} + \frac{a}{\varepsilon} + \frac{\Delta t}{1+\Delta t})M + \kappa A & 0 \\ -\frac{\Delta t}{1+\Delta t} I & I \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_{n+1} \\ \boldsymbol{v}_{n+1} \end{pmatrix} &= \begin{pmatrix} 0 & -\frac{1}{1+\Delta t} M \\ 0 & \frac{1}{1+\Delta t} I \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_n \\ \boldsymbol{v}_n \end{pmatrix} + \begin{pmatrix} G(\boldsymbol{u}_n) \\ 0 \end{pmatrix} \\ &+ \begin{pmatrix} \frac{\sigma}{\sqrt{\Delta t}} M & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{W}^Q_{1,n+1} \\ 0 \end{pmatrix}, \end{aligned}$$

where $G(u_n) = (k(u_n), \psi_i)_{1 \le i \le N_h} \in \mathbb{R}^{N_h}$. As for the parabolic stochastic equation considered in Section 6.3.1, one may expect a numerical strong error for this scheme of order

$$\mathbb{E}(\|(u_t, v_t) - (u_{t,n}, v_{t,n})\|^2)^{\frac{1}{2}} = \mathrm{O}(h + \sqrt{\Delta t}), \tag{6.23}$$

for $\Delta t \le \Delta t_0$. In (6.23), $(u_{t,n}, v_{t,n})_{t \in [0,T]}$ is the interpolation of the discretized point which is piecewise linear in time.
We end up this section with Figure 6.6 which displays simulations of the stochastic Fitzhugh-Nagumo model (6.18) with zero Neumann boundary conditions on a

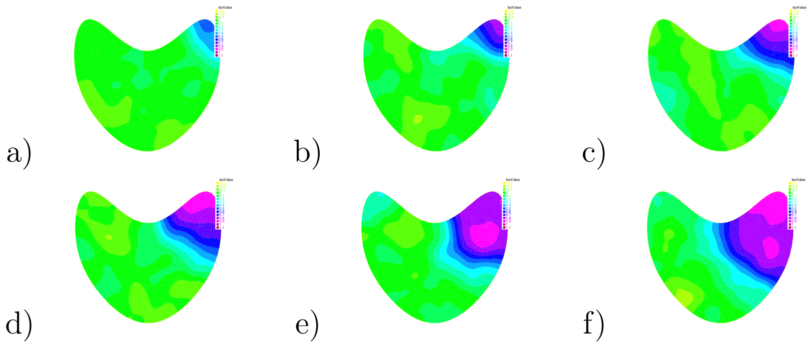


Figure 6.6: Simulations of system (6.18) with $\xi = 2$, $\sigma = 1$, $\varepsilon = 0.1$, $a = 0.1$. These figures must be read from the up-left to the down-right. The time step is $0.05ms$ and there is $0.5ms$ between each figure.

cardioid domain and zero initial conditions. The kernel of the operator Q is given by equation (6.6) for some $\xi > 0$. Due to a strong intensity of the noise source ($\sigma = 1$), we observe the spontaneous nucleation of a front wave with irregular front propagating throughout the whole domain.

6.4 Arrhythmia and reentrant patterns in excitable media

In this section, we focus on classical models for excitable cells, namely Barkley and Mitchell-Schaeffer models. We would like to observe cardiac arrhythmia, that is troubles that may appear in the cardiac beats. Among the diversity of arrhythmia, the phenomena of tachycardia are certainly the most dangerous as they lead to rapid loss of consciousness and death. Tachycardia is described as follows in [JC06]

> The vast majority of tachyarrhythmias are perpetuated by reentrant mechanisms. Reentry occurs when previously activated tissue is repeatedly activated by the propagating action potential wave as it reenters the same anatomical region and reactivates it.

In system (6.1), the equation on u gives the evolution of the cardiac action potential. The equation on v takes into account the evolution of internal biological mechanisms leading to the generation of this action potential. We will be more specifically interested by two systems of this form: the Barkley and Mitchell-Schaeffer models.

6.4.1 Numerical study of the Barkley model

The model

In the deterministic setting, a paradigm for excitable systems where reentrant phenomena such as spiral, meander or scroll waves have been observed and studied is the Barkley model, see [BKT90, Bar91, Bar92, Bar94]. This deterministic model is of the following form

$$\left\{ \begin{array}{lcl} \mathrm{d}u & = & [\kappa \Delta u + \frac{1}{\varepsilon} u(1-u)(u - \frac{v+b}{a})]\mathrm{d}t, \\ \mathrm{d}v & = & [u - v]\mathrm{d}t. \end{array} \right. \tag{6.24}$$

The parameter ε is typically small so that the time scale of u is much faster than that of v. For more details on the dynamic of waves in excitable media, we refer the reader to [Kee80]. The Barkley model, like two-variables models of this type, faithfully captures the behavior of many excitable systems. The deterministic model (6.24) does not exhibit re-entrant patterns unless one imposes special conditions on the domain unless, for instance, one imposes that a portion of the spatial domain is a "dead zone". This means a region with impermeable boundaries where equations (6.24) do not apply: when a wave reaches this dead region, the tip of the wave may turn around and this induces a spiral behavior, see Section 2.2 of [Kee80]. One may also impose specific initial conditions such that some zones are intentionally hyper-polarized: the dead region is somehow transient in this case.

Reentrant patterns

As in [Sha05] we add a colored noise with kernel of type (6.6) to equation (6.24) and so we consider

$$\left\{ \begin{array}{lcl} \mathrm{d}u & = & [\kappa \Delta u + \frac{1}{\varepsilon} u(1-u)(u - \frac{v+b}{a})]\mathrm{d}t + \sigma \mathrm{d}W^{Q_\xi}, \\ \mathrm{d}v & = & [u - v]\mathrm{d}t, \end{array} \right. \tag{6.25}$$

where the kernel of Q_ξ is given by (6.6) for $\xi > 0$.
Figure 6.7 displays a simulation of system (6.25) on the square $D = [0,l] \times [0,l]$ with periodic boundary conditions:

$$\begin{array}{lllll} \forall t \in \mathbb{R}_+, & \forall x \in [0,l] & u_t(x,0) = u_t(x,l), & \text{and} & \frac{\partial u_t}{\partial \vec{n}}(x,0) = \frac{\partial u_t}{\partial \vec{n}}(x,l), \\ & \forall y \in [0,l] & u_t(0,y) = u_t(l,y), & \text{and} & \frac{\partial u_t}{\partial \vec{n}}(0,y) = \frac{\partial u_t}{\partial \vec{n}}(l,y), \end{array} \tag{6.26}$$

where $\vec{n}$ is the external unit normal to the boundary. The numerical scheme is based on the following variational formulation. Given u_0 and v_0 in $H^1(D)$

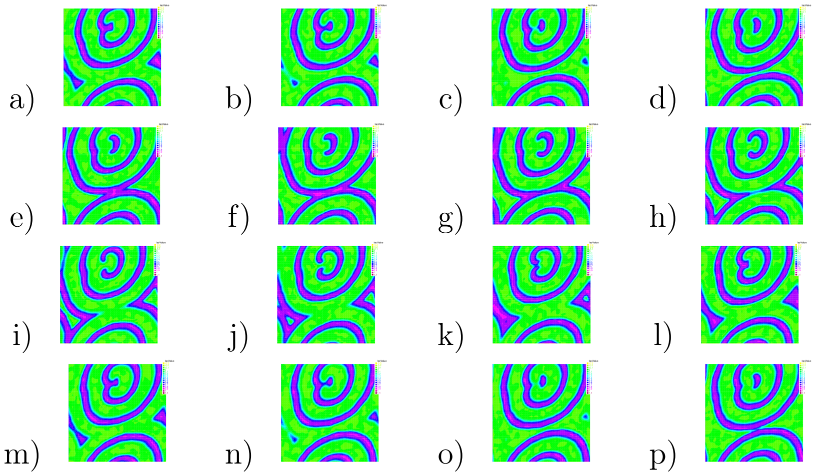


Figure 6.7: Reentry is observed for system (6.25) with $\xi = 2$, $\sigma = 0.15$, $\varepsilon = 0.05$, $a = 0.75$, $b = 0.01$ and $\nu = 1$. These figures must be read from the top-left to the bottom-right. The quiescent state is represented in green whereas the excited state is in violet. If time is recorded in ms, there is $0.5ms$ between each figures for a time step of $0.05ms$.

and satisfying the boundary conditions (6.26), find $(u_n, v_n)_{1\leq n\leq N}$ such that for all $0 \leq n \leq N-1$,

$$\left\{\begin{array}{rcl} (\frac{u_{n+1}-u_n}{\Delta t}, \psi) + \kappa(\nabla u_{n+1}, \nabla\psi) & = & \frac{1}{\varepsilon}(u_n(1-u_n)(u_n - \frac{v_n+b}{a}), \psi) + \frac{\sigma}{\sqrt{\Delta t}}(W^Q_{1,n+1}, \psi), \\ \frac{v_{n+1}-v_n}{\Delta t} & = & u_{n+1} - v_{n+1}. \end{array}\right. \tag{6.27}$$

for all $\psi \in H^1(D)$ satisfying $\psi(x,0) = \psi(x,l)$ and $\psi(0,y) = \psi(l,y)$ for any $(x,y) \in D$. We have solved this problem using the P1 finite element methods, see Section 6.2.2.

Our aim is to observe reentrant patterns generated by the presence of the noise source in this system. Figure 6.7 displays simulations of (6.25) using the P1 finite element method. We observe the spontaneous generation of waves with a reentrant pattern. At some points in the spatial domain, the system is excited and exhibits a reentrant evolution which is self-sustained: a previously activated zone is re-activated by the same wave periodically. As explained in [JC06] and quoted in Section 6.4.1, this phenomenon can be interpreted biologically as tachycardia in the heart tissue. We observe that, as in [Sha05], the constants a and b are chosen such that the deterministic version of system (6.25) may exhibit spiral pattern, see the bifurcation diagram between a and b in [Bar94]. However, in our context, the generation of spiral is a phenomenon which is due solely to the presence of noise. In particular, there is no need for a "dead region", as previously mentioned for the

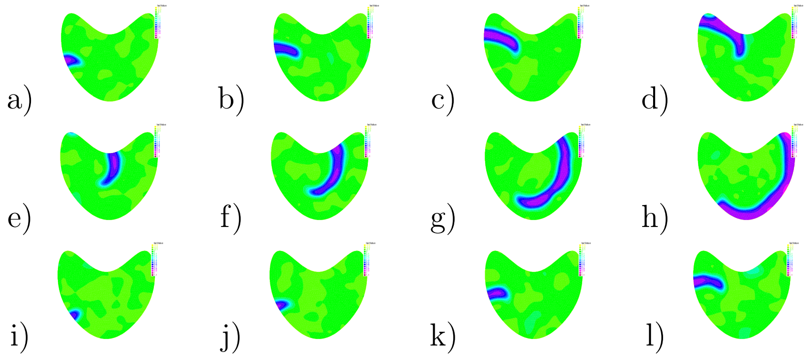


Figure 6.8: Simulations of system (6.25) with $\xi = 2$, $\sigma = 0.15$, $\varepsilon = 0.05$, $a = 0.75$, $b = 0.01$ and $\nu = 1$. As for the previous figure, the quiescent state is represented in green whereas the excited state is in violet. Another phenomena of re-entry is observed on this cardioid geometry with zero Neumann boundary conditions. There is $2ms$ between each snapshot for a time step for the simulations equals to $0.05ms$.

observation of spirals or reentrant patterns in a deterministic context. Figure 6.8 displays a simulation of system (6.25) on a cardioid domain with zero Neumann boundary conditions, see (6.19). We observe the spontaneous generation of a wave turning around itself like a spiral and thus reactivating zones already activated by the same wave.

To gain a better insight into these reentrant phenomena, a bifurcation diagram between ε and σ in system (6.25) is displayed in Figure 6.9. In this figure, the other parameters a, b, ν, ξ are held fixed. The domain and boundary conditions are the same as for Figure 6.7. Three distinct areas emerge from repeated simulations:

- the area NW (for No Wave) where no wave is observed.
- the area W (for Wave) where at least one wave is generated on average. Such waves do not exhibit reentrant patterns.
- the area RW (for Reentrant Wave) where waves with re-entry are observed. The wave has the same pattern as in Figure 6.7.

At transition between the areas W and RW, ring waves with the same pattern as reentrant waves may be observed: two arms which join each other to form a ring. We also remark that for a fixed ε, when σ increases, the number of nucleated waves increases. On the contrary, for a fixed σ when ε increases the number of nucleated waves decreases. Let us notice that for small ε, that is when the transition between the quiescent and excited state is very sharp, small noise may powerfully initiate

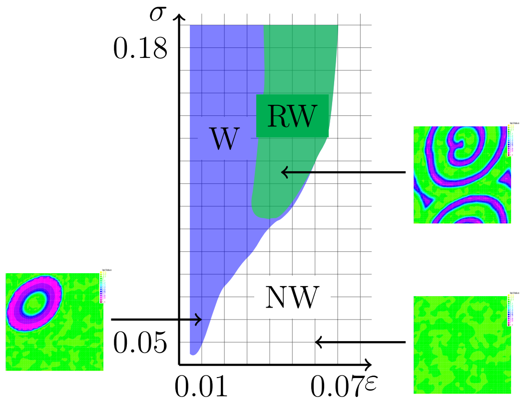


Figure 6.9: Numerical bifurcation diagram between ε and σ of system (6.25) with $\xi = 2$, $a = 0.75$, $b = 0.01$ and $\nu = 1$ held fixed.

spike. However, we only observe reentrant patterns when ε is large enough. Notice also that the separation curve between the zone NW and the two zones W and RW is exponentially shaped. This may be related to the large deviation theory for slow-fast system of SPDE.

6.4.2 Numerical study of the Mitchel-Schaeffer model

The model

Fitzhugh-Nagumo model is the most popular phenomenological model for cardiac cells. However this model has some flows, in particular the hyperpolarization and the stiff slope in the repolarization phase. The Mitchell-Schaeffer model [MS03] has been proposed to improve the shape of the action potential in cardiac cells. The spatial version of the Mitchell-Schaeffer model reads as follows

$$\left\{\begin{array}{lcl} \mathrm{d}u & = & \left[\kappa\Delta u + \dfrac{v}{\tau_{\text{in}}} u^2(u-1) + \dfrac{u}{\tau_{\text{out}}}\right] \mathrm{d}t + \sigma \mathrm{d}W^{Q_\xi}, \\ \mathrm{d}v & = & \left[\dfrac{1}{\tau_{\text{open}}}(v-1)\, 1_{u<u_{\text{gate}}} + \dfrac{v}{\tau_{\text{close}}} 1_{u \geq u_{\text{gate}}}\right] \mathrm{d}t. \end{array}\right. \tag{6.28}$$

The numerical scheme is based on the following variational formulation. Given u_0 and v_0 in $H^1(D)$, find (u_n, v_n) such that for all $0 \leq n \leq N-1$,

$$\left\{\begin{array}{lcl} (\frac{u_{n+1}-u_n}{\Delta t}, \psi) + \kappa(\nabla u_{n+1}, \nabla\psi) & = & \frac{1}{\varepsilon}(\frac{v_n}{\tau_{\text{in}}} u_n^2(u_n-1) + \frac{1}{\tau_{\text{out}}} u_n, \psi) + \frac{\sigma}{\sqrt{\Delta t}}(W_{n+1}^Q(1), \psi), \\ \frac{v_{n+1}-v_n}{\Delta t} & = & \dfrac{1}{\tau_{\text{open}}}(v_n - 1)\, 1_{u_n<u_{\text{gate}}} + \dfrac{v_n}{\tau_{\text{close}}} 1_{u_n \geq u_{\text{gate}}} \end{array}\right. \tag{6.29}$$

for $\psi \in H^1(D)$. More precisely, we solve this problem with the P1 finite element method.

Numerical investigations

Bifurcations have been investigated in Figure 6.10 for the same domain and boundary conditions as for the bifurcation diagram related to Barkley model (Figure 6.9). We choose to fix all the parameters except the intensity of the noise σ and τ_{close} to investigate the influence of the strength of the noise and the characteristic time for the recovery variable v to get closed. From repeated simulations, five distinct areas emerge:

- the area NW (for No Wave) where no wave is observed.
- the area W (for Wave) where at least one wave is generated on average. These waves do not exhibit reentrant patterns. However, these waves may be generated with the same pattern as reentrant waves: two arms which meet up and agree to form a ring.
- the area RW (for Reentrant Wave) where waves with re-entry may be observed as in Figure 6.7.
- the area DW (for Disorganized Wave) where reentrant waves are initiated but break down in numerous pieces resulting in a very disorganized evolution. In a sense, this disorganized evolution may be regarded as reentrant since previously activated zone may be re-activated by one of these resulting pieces.
- the area T (for Transition) is a transition area between reentrant waves and more disorganized patterns as observed in the area DW.

Appendix 6.A Proof of Theorem 6.2.1

Recall that the domain D is polyhedral such that

$$\overline{D} = \bigcup_{T \in \mathcal{T}_h} T.$$

Let $i \in \{0, 0_a, 1\}$. The process $(D_h(t), t \in [0, \tau])$ defined by

$$D_h(t) = W_t^Q - W_t^{Q,h,i}$$

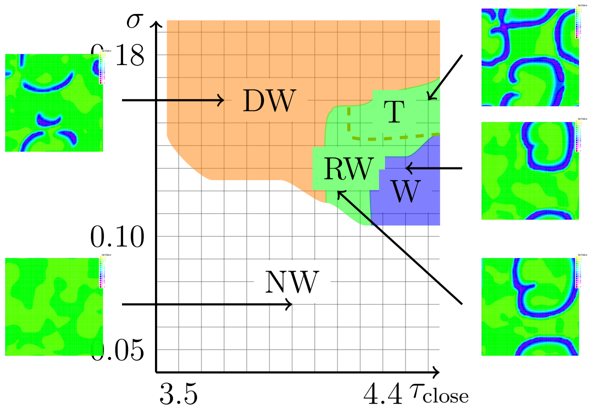


Figure 6.10: Numerical bifurcation diagram between τ_{close} and σ of system (6.28) with $\xi = 2$, $\tau_{\text{in}} = 0.07$, $\tau_{\text{out}} = 0.7$, $\tau_{\text{open}} = 8.$, $u_{\text{gate}} = 0.13$ and $\nu = 0.03$ held fixed.

is a centered Wiener process. In particular, it is a continuous martingale and thus, by the Burkholder-Davis-Gundy inequality (see Theorem 3.4.9 of [PZ07]) we have

$$\mathbb{E}\left(\sup_{t\in[0,\tau]} \|D_h(t)\|^2\right) \leq c_2 \mathbb{E}(\|D_h(\tau)\|^2)$$

with c_2 a constant which does not depend on h or τ. We begin with the case $i = 1$. Since the processes W^Q and $W^{Q,h,1}$ are regular in space, we write

$$\mathbb{E}(\|D_h(\tau)\|^2) = \mathbb{E}\left(\int_D (W^Q_\tau(x) - W^{Q,h,1}_\tau(x))^2 \mathrm{d}x\right).$$

We use the definition of $W^{Q,h,1}$ in Definition 6.2.3 and the fact that $\sum_{i=1}^{N_h} \psi_i = 1$ to obtain

$$\begin{aligned}
\mathbb{E}(\|D_h(\tau)\|^2) &= \mathbb{E}\left(\int_D (W^Q_\tau(x) - \sum_{i=1}^{N_h} W^Q_\tau(P_i)\psi_i(x))^2 \mathrm{d}x\right) \\
&= \mathbb{E}\left(\int_D (\sum_{i=1}^{N_h} (W^Q_\tau(x) - W^Q_\tau(P_i))\psi_i(x))^2 \mathrm{d}x\right) \\
&= \mathbb{E}\left(\int_D \sum_{i,j=1}^{N_h} (W^Q_\tau(x) - W^Q_\tau(P_i))(W^Q_\tau(x) - W^Q_\tau(P_j))\psi_i(x)\psi_j(x)\mathrm{d}x\right).
\end{aligned}$$

By an application of Fubini's theorem, exchanging over the expectation, integral and summation, we get

$$\begin{aligned}\mathbb{E}(\|D_h(\tau)\|^2) &= \sum_{i,j=1}^{N_h}\int_D \mathbb{E}\left((W_\tau^Q(x)-W_\tau^Q(P_i))(W_\tau^Q(x)-W_\tau^Q(P_j))\right)\psi_i(x)\psi_j(x)\mathrm{d}x\\ &= \tau\sum_{i,j=1}^{N_h}\int_D (C(0)-C(P_i-x)-C(P_j-x)+C(P_i-P_j))\,\psi_i(x)\psi_j(x)\mathrm{d}x.\end{aligned}$$

For all $1\le i,j\le N_h$, if the intersection of the supports of ψ_i and ψ_j is not empty, then

$$\forall x\in \mathrm{supp}\psi_i, \forall y\in\mathrm{supp}\psi_j, |x-y|\le Kh.$$

Thus, there exists $K>0$ such that, for all i,j, if $\mathrm{supp}\psi_i\cap\mathrm{supp}\psi_j\neq\emptyset$ and $x\in\mathrm{supp}\psi_i\cap\mathrm{supp}\psi_j$, a Taylor's expansion yields

$$|C(0)-C(P_i-x)-C(P_j-x)+C(P_i-P_j)|\le Kh^2,$$

where we have used the fact that $\nabla C(0)=0$. Then,

$$\mathbb{E}(\|D_h(\tau)\|^2)\le K\tau h^2.$$

This ends the proof for the case $i=1$. The case $i=0$ can be treated similarly. For the case $i=0_a$, we proceed as follows. Since the processes W^Q and $W^{Q,h,0_a}$ are regular in space, as before we write

$$\mathbb{E}(\|D_h(\tau)\|^2)=\mathbb{E}\left(\int_D (W_\tau^Q(x)-W_\tau^{Q,h,0_a}(x))^2\mathrm{d}x\right).$$

We use the definition of $W^{Q,h,0_a}$ and develop the square to obtain

$$\begin{aligned}\mathbb{E}(\|D_h(\tau)\|^2)=\mathbb{E}\Bigg(\int_D W_\tau^Q(x)^2 &- 2\sum_{T\in\mathcal{T}_h}\frac{1}{|T|}(W_\tau^Q,1_T)W_\tau^Q(x)1_T(x)\\ &+\sum_{T,S\in\mathcal{T}_h}\frac{1}{|T||S|}(W_\tau^Q,1_T)(W_\tau^Q,1_S)1_T(x)1_S(x)\mathrm{d}x\Bigg).\end{aligned}$$

To simplify the above expression, we use the fact that the triangles do not intersect to obtain

$$\begin{aligned}\mathbb{E}(\|D_h(\tau)\|^2) &= \mathbb{E}\left(\int_D W_\tau^Q(x)^2 - 2\sum_{T\in\mathcal{T}_h}\frac{1}{|T|}(W_\tau^Q, 1_T)W_\tau^Q(x)1_T(x)\right.\\ &\quad\left. + \sum_{T\in\mathcal{T}_h}\frac{1}{|T|^2}(W_\tau^Q, 1_T)^2 1_T(x)\mathrm{d}x\right)\\ &= \mathbb{E}\left(\int_D W_\tau^Q(x)^2\mathrm{d}x - \sum_{T\in\mathcal{T}_h}\frac{1}{|T|}(W_\tau^Q, 1_T)^2\right).\end{aligned}$$

By an application of Fubini's theorem, exchanging over the expectation and summation, we get

$$\mathbb{E}(\|D_h(\tau)\|^2) = \tau\left(C(0)|D| - \sum_{T\in\mathcal{T}_h}\frac{1}{|T|}(Q1_T, 1_T)\right). \tag{6.30}$$

Since $\overline{D} = \bigcup_{T\in\mathcal{T}_h} T$ we have

$$C(0)|D| = \sum_{T\in\mathcal{T}_h}\frac{1}{|T|}\int_T\int_T C(0)\mathrm{d}z_1\mathrm{d}z_2,$$

hence, plugging in (6.30)

$$\mathbb{E}(\|D_h(\tau)\|^2) = \tau\sum_{T\in\mathcal{T}_h}\frac{1}{|T|}\int_T\int_T[C(0) - C(z_1 - z_2)]\mathrm{d}z_1\mathrm{d}z_2. \tag{6.31}$$

Thanks to the fact that $\nabla C = 0$, a Taylor's expansion yields

$$C(0) - C(z_1 - z_2) = (z_1 - z_2)\cdot \mathrm{Hess}C(0)(z_1 - z_2) + \mathrm{o}(|z_1 - z_2|^2) \tag{6.32}$$

Thus, thanks to (6.2), for all z_1, z_2 in the same triangle T

$$|C(0) - C(z_1 - z_2)| \le Kh^2,$$

where K is independent from $T \in \mathcal{T}_h$. Plugging in (6.31) yields

$$\mathbb{E}(\|D_h(\tau)\|^2) \le K\tau h^2$$

for a deterministic constant K.

Appendix 6.B Proof of Theorem 6.3.1

Let $1 \leq i \leq I_h$ and $n \in \{0, \ldots, N-1\}$. A direct calculation using (6.9) and (6.10) gives

$$\frac{1}{\Delta t}(e^h_{n+1} - e^h_n, w_{i,h}) + a(e^h_{n+1}, w_{i,h}) = (\tilde{\varepsilon}^h_n, w_{i,h})$$

with

$$(\tilde{\varepsilon}^h_n, w_{i,h}) = \frac{\sigma}{\Delta t}(W^{Q,h}_{n+1} - W^{Q,h}_n, w_{i,h}) - \frac{1}{\Delta t}(\Pi_h u_{t_{n+1}} - \Pi_h u_{t_n}, w_{i,h}) - a(u_{t_{n+1}}, w_{i,h}).$$

Then, using the fact that $a(e^h_{n+1}, w_{i,h}) = \lambda_{i,h}(e^h_{n+1}, w_{i,h})$ we obtain

$$(e^h_{n+1}, w_{i,h}) = \frac{1}{1+\Delta t \lambda_{i,h}}(e^h_n, w_{i,h}) + \frac{\Delta t}{1+\Delta t \lambda_{i,h}}(\tilde{\varepsilon}^h_n, w_{i,h}). \tag{6.33}$$

Moreover, we set $\mu_{i,h,\Delta t} = \frac{\log(1+\Delta t \lambda_{i,h})}{\Delta t} > 0$ such that

$$e^{-\mu_{i,h,\Delta t}\Delta t} = \frac{1}{1+\Delta t \lambda_{i,h}}.$$

By induction from (6.33) we obtain, for all $1 \leq i \leq I_h$ and $0 \leq n \leq N-1$:

$$(e^h_n, w_{i,h}) = e^{-\mu_{i,h,\Delta t} n \Delta t}(e^h_0, w_{i,h}) + \Delta t \sum_{k=0}^{n-1} e^{-\mu_{i,h,\Delta t}(n-k)\Delta t}(\tilde{\varepsilon}^h_k, w_{i,h}).$$

We multiply by $w_{i,h}$ and sum over i to obtain the desired result (6.12) with $\tilde{\varepsilon}$ instead of ε:

$$e^h_n = \sum_{i=1}^{I_h} e^{-\mu_{i,h,\Delta t} n \Delta t}(e^h_0, w_{i,h}) w_{i,h} + \Delta t \sum_{i=1}^{I_h} \sum_{k=0}^{n-1} e^{-\mu_{i,h,\Delta t}(n-k)\Delta t}(\tilde{\varepsilon}^h_k, w_{i,h}) w_{i,h}. \tag{6.34}$$

Then, we notice that, by Itô formula from the variational formulation satisfied by u:

$$(u_{(n+1)\Delta t} - u_{n\Delta t}, w_{i,h}) = -\int_{n\Delta t}^{(n+1)\Delta t} a(u_s, w_{i,h}) \mathrm{d}s + \sigma(W^Q_{(n+1)\Delta t} - W^Q_{n\Delta t}, w_{i,h}).$$

Now we add and substract the term $\frac{1}{\Delta t}\int_{n\Delta t}^{(n+1)\Delta t} a(u_s, w_{i,h})\mathrm{d}s$ in the definition of $(\tilde{\varepsilon}_k^h, w_{i,h})$ and use the above Itô formula to get:

$$
\begin{aligned}
(\tilde{\varepsilon}_n^h, w_{i,h}) &= \frac{\sigma}{\Delta t}(W_{n+1}^{Q,h} - W_n^{Q,h}, w_{i,h}) - \frac{1}{\Delta t}(\Pi_h u_{t_{n+1}} - \Pi_h u_{t_n}, w_{i,h}) - \frac{1}{\Delta t}\int_{n\Delta t}^{(n+1)\Delta t} a(u_s, w_{i,h})\mathrm{d}s \\
&+ \frac{1}{\Delta t}\int_{n\Delta t}^{(n+1)\Delta t} a(u_s, w_{i,h})\mathrm{d}s - a(u_{t_{n+1}}, w_{i,h}) \\
&= \frac{\sigma}{\Delta t}(W_{n+1}^{Q,h} - W_n^{Q,h}, w_{i,h}) - \frac{\sigma}{\Delta t}(W_{(n+1)\Delta t}^{Q} - W_{n\Delta t}^{Q}, w_{i,h}) \\
&+ \frac{1}{\Delta t}\int_{n\Delta t}^{(n+1)\Delta t} a(u_s, w_{i,h})\mathrm{d}s - a(u_{\Delta t(n+1)}, w_{i,h}) + \frac{1}{\Delta t}((\mathrm{I} - \Pi_h)(u_{(n+1)\Delta t} - u_{n\Delta t}), w_{i,h}) \\
&= (\varepsilon_n^h, w_{i,h}).
\end{aligned}
$$

Plugging this equality in (6.34), we get the result.

Appendix 6.C Proofs of Lemmas 6.3.1 and 6.3.2

We start with a technical lemma.

Lemma 6.C.1. *Let us define, for $i, n \geq 1$, $h, \Delta t > 0$,*

$$
\alpha_{i,n,h,\Delta t} = \sum_{k=0}^{n-1} \int_{k\Delta t}^{(k+1)\Delta t} \left(e^{-\mu_{i,h,\Delta t}(n\Delta t - k\Delta t)} - e^{-\mu_{i,h,\Delta t}(n\Delta t - s)}\right)^2 \mathrm{d}s.
$$

Then

$$
\alpha_{i,n,h,\Delta t} \leq \Delta t f(\Delta t \lambda_{i,h}), \tag{6.35}
$$

where the function f is defined on $\mathbb{R}_+^*$ by

$$
f(x) = \frac{\log(1+x) + \frac{1}{2}x^2 - x}{x(x+2)\log(1+x)}.
$$

Proof. Since i and h are held fixed, we simply write in this proof $\mu_{\Delta t}$ and λ for

$\mu_{i,h,\Delta t}$ and $\lambda_{i,h}$. We have

$$\begin{aligned}
&\sum_{k=0}^{n-1}\int_{k\Delta t}^{(k+1)\Delta t}\left(e^{-\mu_{\Delta t}(n\Delta t-k\Delta t)}-e^{-\mu_{\Delta t}(n\Delta t-s)}\right)^2\mathrm{d}s\\
=&\sum_{k=0}^{n-1}e^{-2\mu_{\Delta t}(n\Delta t-k\Delta t)}\int_{k\Delta t}^{(k+1)\Delta t}\left(1-e^{-\mu_{\Delta t}(k\Delta t-s)}\right)^2\mathrm{d}s\\
=&\sum_{k=0}^{n-1}e^{-2\mu_{\Delta t}(n-k)\Delta t}\frac{1}{\mu_{\Delta t}}\int_0^{\mu_{\Delta t}\Delta t}(1-e^s)^2\,\mathrm{d}s\\
=&\frac{1}{\mu_{\Delta t}}\int_0^{\mu_{\Delta t}\Delta t}(1-e^s)^2\,\mathrm{d}s\frac{1-e^{-2\mu_{\Delta t}n\Delta t}}{1-e^{-2\mu_{\Delta t}\Delta t}}e^{-2\mu_{\Delta t}\Delta t}\\
=&\frac{1}{\mu_{\Delta t}}\left[\mu_{\Delta t}\Delta t-2e^{\mu_{\Delta t}\Delta t}+\frac{1}{2}e^{2\mu_{\Delta t}\Delta t}+2-\frac{1}{2}\right]\frac{1-e^{-2\mu_{\Delta t}n\Delta t}}{e^{2\mu_{\Delta t}\Delta t}-1}.
\end{aligned}$$

Using the definition of $\mu_{\Delta t}$ this leads to

$$\begin{aligned}
&\sum_{k=0}^{n-1}\int_{k\Delta t}^{(k+1)\Delta t}\left(e^{-\mu_{\Delta t}(n\Delta t-k\Delta t)}-e^{-\mu_{\Delta t}(n\Delta t-s)}\right)^2\mathrm{d}s\\
=&\frac{\Delta t}{\log(1+\Delta t\lambda)}\left[\log(1+\Delta t\lambda)-2(1+\Delta t\lambda)+\frac{1}{2}(1+\Delta t\lambda)^2+2-\frac{1}{2}\right]\frac{1-e^{-2\mu_{\Delta t}n\Delta t}}{(1+\Delta t\lambda)^2-1}\\
\leq&\frac{\Delta t}{\log(1+\Delta t\lambda)}\left[\log(1+\Delta t\lambda)+\frac{1}{2}(\Delta t\lambda)^2-\Delta t\lambda\right]\frac{1}{(1+\Delta t\lambda)^2-1}\\
=&\Delta t f(\Delta t\lambda),
\end{aligned}$$

as required. □

Proof of Lemma 6.3.1. We start with the detailed proof for the estimation of the quantity $\mathbb{E}\left(\|A_n-B_n\|^2\right)$. The estimation of $\mathbb{E}\left(\|C_n-D_n\|^2\right)$ is obtained the same way. Let us write

$$\sum_{k=0}^{n-1}e^{-\mu_{i,h,\Delta t}(n\Delta t-k\Delta t)}(W^Q_{(k+1)\Delta t}-W^Q_{k\Delta t},w_{i,h})=\sum_{k=0}^{n-1}\left(\int_{k\Delta t}^{(k+1)\Delta t}e^{-\mu_{i,h,\Delta t}(n\Delta t-k\Delta t)}\mathrm{d}W^Q_s,w_{i,h}\right)$$

and

$$\left(\int_0^{\Delta t n}e^{-\mu_{i,h,\Delta t}(n\Delta t-s)}\mathrm{d}W^Q_s,w_{i,h}\right)=\sum_{k=0}^{n-1}\left(\int_{k\Delta t}^{(k+1)\Delta t}e^{-\mu_{i,h,\Delta t}(n\Delta t-s)}\mathrm{d}W^Q_s,w_{i,h}\right).$$

Then A_n-B_n is given by

$$\sum_{i=1}^{I_h}\sum_{k=0}^{n-1}\left(\int_{k\Delta t}^{(k+1)\Delta t}e^{-\mu_{i,h,\Delta t}(n\Delta t-k\Delta t)}-e^{-\mu_{i,h,\Delta t}(n\Delta t-s)}\mathrm{d}W^Q_s,w_{i,h}\right)w_{i,h}.$$

Since the family $(w_{i,h}, 1 \leq i \leq I_h)$ is orthonormal in H we have

$$\begin{aligned}
&\mathbb{E}(\|A_n - B_n\|^2) \\
&= \sum_{i=1}^{I_h} \mathbb{E}\left(\sum_{k=0}^{n-1}\left(\int_{k\Delta t}^{(k+1)\Delta t} e^{-\mu_{i,h,\Delta t}(n\Delta t - k\Delta t)} - e^{-\mu_{i,h,\Delta t}(n\Delta t - s)} \mathrm{d}W_s^Q, w_{i,h}\right)\right)^2.
\end{aligned}$$

Completing the basis $(w_{i,h}, 1 \leq i \leq I_h)$ of V_h, which is orthonormal in H, into a Hilbert basis $(w_{i,h}, i \geq 1)$ of H, one can write

$$W_s^Q = \sum_{j \geq 1} \beta_{j,h,s} Q^{\frac{1}{2}} w_{j,h},$$

where $(\beta_{j,h,s}, j \geq 1)$ is a sequence of independent brownian motions. This gives

$$\begin{aligned}
&\mathbb{E}(\|A_n - B_n\|^2) \\
&= \sum_{i=1}^{I_h} \mathbb{E}\left(\sum_{j \geq 1}\sum_{k=0}^{n-1}\int_{\Delta t k}^{\Delta t(k+1)} e^{-\mu_{i,h,\Delta t}(n\Delta t - k\Delta t)} - e^{-\mu_{i,h,\Delta t}(n\Delta t - s)} \mathrm{d}\beta_{j,h,s}(Q^{\frac{1}{2}} w_{j,h}, w_{i,h})\right)^2.
\end{aligned} \tag{6.36}$$

Remember that the brownian motions $(\beta_{j,h,s}, j \geq 1)$ are independent and remark that for each $j \geq 1$, the sequence of variables

$$\left(\int_{k\Delta t}^{(k+1)\Delta t} e^{-\mu_{i,h,\Delta t}(n\Delta t - k\Delta t)} - e^{-\mu_{i,h,\Delta t}(n\Delta t - s)} \mathrm{d}\beta_{j,h,s}\right)_{0 \leq k \leq n-1}$$

is a sequence of independent real-valued random variables. Hence, by independence and Itô isometry

$$\begin{aligned}
&\mathbb{E}\left(\sum_{j \geq 1}\sum_{k=0}^{n-1}\int_{k\Delta t}^{(k+1)\Delta t} e^{-\mu_{i,h,\Delta t}(n\Delta t - k\Delta t)} - e^{-\mu_{i,h,\Delta t}(n\Delta t - s)} \mathrm{d}\beta_{j,h,s}(Q^{\frac{1}{2}} w_{j,h}, w_{i,h})\right)^2 \\
&= \sum_{j \geq 1}(Q^{\frac{1}{2}} w_{j,h}, w_{i,h})^2 \sum_{k=0}^{n-1}\int_{k\Delta t}^{(k+1)\Delta t}\left(e^{-\mu_{i,h,\Delta t}(n\Delta t - k\Delta t)} - e^{-\mu_{i,h,\Delta t}(n\Delta t - s)}\right)^2 \mathrm{d}s \\
&= \sum_{j \geq 1}(Q^{\frac{1}{2}} w_{j,h}, w_{i,h})^2 \alpha_{i,n,h,\Delta t},
\end{aligned}$$

where $\alpha_{i,n,h,\Delta t}$ is defined in Lemma 6.C.1. Plugging this identity in (6.36) gives

$$\mathbb{E}(\|A_n - B_n\|^2) = \sum_{j \geq 1}\sum_{i=1}^{I_h}(Q^{\frac{1}{2}} w_{j,h}, w_{i,h})^2 \alpha_{i,n,h,\Delta t}. \tag{6.37}$$

Using the estimation of $\alpha_{i,n,h,\Delta t}$ from Lemma 6.C.1 leads to

$$\mathbb{E}(\|A_n - B_n\|^2) \leq \sum_{j\geq 1}\sum_{i=1}^{I_h}(Q^{\frac{1}{2}}w_{j,h}, w_{i,h})^2\Delta t f(\Delta t\lambda_{i,h})$$

Notice that the function f is positive and bounded by a constant K_f on $\mathbb{R}_+$. Therefore

$$\mathbb{E}(\|A_n - B_n\|^2) \leq K_f\Delta t\sum_{j\geq 1}\sum_{i=1}^{I_h}(Q^{\frac{1}{2}}w_{j,h}, w_{i,h})^2.$$

Finally, since $\sum_{j\geq 1}\sum_{i\geq 1}(Q^{\frac{1}{2}}w_{j,h}, w_{i,h})^2 = \mathrm{Tr}(Q)$, we obtain

$$\mathbb{E}(\|A_n - B_n\|^2) \leq K_f\mathrm{Tr}(Q)\Delta t.$$

We go on with the estimation of the quantity $\mathbb{E}(\|B_n - D_n\|^2)$. Notice that we can write

$$B_n - D_n = \sum_{i=1}^{I_h}\left(\int_0^{n\Delta t} e^{-\mu_{i,h,\Delta t}(n\Delta t - s)}\mathrm{d}(W_s^Q - W_s^{Q,h,1}), w_{i,h}\right)w_{i,h}.$$

Proceeding as above

$$\mathbb{E}(\|B_n - D_n\|^2) = \sum_{j\geq 1}\sum_{i=1}^{I_h}((Q^{\frac{1}{2}} - (Q^{h,1})^{\frac{1}{2}})w_{j,h}, w_{i,h})^2\int_0^{n\Delta t} e^{-2\mu_{i,h,\Delta t}(n\Delta t - s)}\mathrm{d}s.$$

Note that $\int_0^{n\Delta t} e^{-2\mu_{i,h,\Delta t}(n\Delta t-s)}\mathrm{d}s \leq \frac{1}{2\mu_{i,h,\Delta t}}$. We get

$$\begin{aligned}\mathbb{E}(\|B_n - D_n\|^2) &\leq \sum_{j\geq 1}\sum_{i=1}^{I_h}\frac{((Q^{\frac{1}{2}} - (Q^{h,1})^{\frac{1}{2}})w_{j,h}, w_{i,h})^2}{2\mu_{i,h,\Delta t}}\\ &= \sum_{j\geq 1}\sum_{i=1}^{I_h}((Q^{\frac{1}{2}} - (Q^{h,1})^{\frac{1}{2}})w_{j,h}, w_{i,h})^2\frac{\Delta t}{2\log(1+\Delta t\lambda_{i,h})}.\end{aligned}$$

For any $\Delta t_0 > 0$ we have

$$\sup_{\Delta t\in[0,\Delta t_0]}\frac{\Delta t}{2\log(1+\Delta t\lambda_{i,h})} \leq K_1 = \sup_{\Delta t\in[0,\Delta t_0]}\frac{\Delta t}{2\log(1+\Delta t\lambda_1)} < \infty,$$

we obtain

$$\mathbb{E}(\|B_n - D_n\|^2) \leq K_1\sum_{i,j\geq 1}((Q^{\frac{1}{2}} - (Q^{h,1})^{\frac{1}{2}})w_{j,h}, w_{i,h})^2 = K_1\mathbb{E}(\|W_1^Q - W_1^{Q,h}\|^2).$$

By Theorem 6.2.1, we obtain that there exists a constant K depending only on $|D|$ such that

$$\mathbb{E}(\|B_n - D_n\|^2) \le Kh^2.$$

□

Proof of Lemma 6.3.2. The error E_n, for $1 \le n \le N$ splits itself into the three following terms:

$$E_n = \alpha_n + \beta_n + \gamma_n$$

where

$$\alpha_n = \sum_{i=1}^{I_h} \sum_{k=0}^{n-1} e^{-\mu_{i,h,\Delta t}(n\Delta t - k\Delta t)}((\mathrm{I} - \Pi_h)e^{A\Delta tk}(e^{A\Delta t} - \mathrm{I})\zeta, w_{i,h})w_{i,h},$$

$$\beta_n = \sum_{i=1}^{I_h} \sum_{k=0}^{n-1} e^{-\mu_{i,h,\Delta t}(n\Delta t - k\Delta t)}((\mathrm{I} - \Pi_h)\int_0^{\Delta tk} e^{A(\Delta tk - s)}(e^{A\Delta t} - \mathrm{I})\mathrm{d}W_s^Q, w_{i,h})w_{i,h},$$

$$\gamma_n = \sum_{i=1}^{I_h} \sum_{k=0}^{n-1} e^{-\mu_{i,h,\Delta t}(n\Delta t - k\Delta t)}((\mathrm{I} - \Pi_h)\int_{\Delta tk}^{\Delta t(k+1)} e^{A(\Delta t(k+1) - s)}\mathrm{d}W_s^Q, w_{i,h})w_{i,h}.$$

We begin with the estimation of $\mathbb{E}(\|\alpha_n\|^2)$. Using elementary calculations we obtain

$$\alpha_n = (\mathrm{I} - \Pi_h)S_n\zeta$$

with the operator S_n defined by

$$S_n = (e^{A\Delta t} - \mathrm{I}))(\mathrm{I} + A\Delta t)^{-n}(\mathrm{I} - e^{A\Delta t}(\mathrm{I} + A\Delta t))^{-1}(\mathrm{I} - e^{An\Delta t}(\mathrm{I} + A\Delta t)^n)$$

Using the spectral decomposition of A, it is not difficult to show that the operator S_n is bounded uniformly in n and Δt. Thus, using Lemma 6.5.1 of [RT83] to control $\mathrm{I} - \Pi_h$, we obtain that there exists a constant C (independent of n) such that

$$\mathbb{E}(\|\alpha_n\|^2) \le Ch^2 \sum_{p=1}^{2} \mathbb{E}(\|\nabla^p \zeta\|^2).$$

Recall that $\zeta \in H^2(D)$ such that the above quantity is finite. Let us pursue with the estimation on γ_n. Using similar calculation as in Lemma 6.3.1 we obtain

$$\mathbb{E}(\|\gamma_n\|^2) = \sum_{i=1}^{I_h} \sum_{k=0}^{n-1} \sum_{j\ge 1} e^{-2\mu_{i,h,\Delta t}(\Delta tn - \Delta tk)} \int_{\Delta tk}^{\Delta t(k+1)} e^{-2\lambda_{i,h}(\Delta t(k+1) - s)}\mathrm{d}s((\mathrm{I} - \Pi_h)Q^{\frac{1}{2}}w_{j,h}, w_{i,h})^2.$$

And then

$$\begin{aligned}\mathbb{E}(\|\gamma_n\|^2) &= \sum_{i=1}^{I_h}\sum_{j\geq 1}\frac{1-e^{-2\lambda_{j,h}\Delta t}}{2\lambda_{j,h}}((\mathrm{I}-\Pi_h)Q^{\frac{1}{2}}w_{j,h},w_{i,h})^2\frac{1}{\Delta t\lambda_{i,h}}\\ &\leq \frac{1}{\lambda_1}\sup_{x\geq 0}\left(\frac{1-e^{-2x}}{2x}\right)\mathrm{Tr}((\mathrm{I}-\Pi_h)Q(\mathrm{I}-\Pi_h)^*)\leq Ch^2.\end{aligned}$$

for some constant C. Notice that we have used the fact that there exists a constant C such that:

$$\mathrm{Tr}((\mathrm{I}-\Pi_h)Q(\mathrm{I}-\Pi_h)^*)\leq Ch^2\sum_{p=1}^{2}\mathrm{Tr}(\nabla^p Q).$$

The above quantity is finite since $\mathrm{Tr}(\Delta Q)<\infty$ in our setting. The estimates on $\mathbb{E}(\|\beta_n\|^2)$ and $\mathbb{E}(\|F_n\|^2)$ are obtained using similar arguments. $\square$

Bibliography

[ABG$^+$13] R. Azaïs, J-B. Bardet, A. Genadot, N. Krell, and P-A. Zitt, *Piecewise Deterministic Markov Processes (PDMPs). Recent Results*, Submitted (2013).

[ADGP12a] R. Azaïs, F. Dufour, and A. Gégout-Petit, *Nonparametric estimation of the conditional distribution of the inter-jumping times for piecewise-deterministic Markov processes.*

[ADGP12b] ———, *Nonparametric estimation of the jump rate for non-homogeneous marked renewal processes*, arXiv preprint arXiv:1202.2211 (2012).

[Ald78] D. Aldous, *Stopping times and tightness*, The Annals of Probability (1978), 335–340.

[ANZ98] E. Allen, S. Novosel, and Z. Zhang, *Finite element and difference approximation of some linear stochastic partial differential equations*, Stochastics: An International Journal of Probability and Stochastic Processes **64** (1998), no. 1-2, 117–142.

[Aus08] T. Austin, *The emergence of the deterministic Hodgkin-Huxley equations as a limit from the underlying stochastic ion-channel mechanism*, The Annals of Applied Probability **18** (2008), no. 4, 1279–1325.

[Aza12] R. Azaïs, *A recursive nonparametric estimator for the transition kernel of a piecewise-deterministic Markov process*, arXiv preprint arXiv:1211.5579 (2012).

[Bar91] D. Barkley, *A model for fast computer simulation of waves in excitable media*, Physica D: Nonlinear Phenomena **49** (1991), no. 1-2, 61–70.

[Bar92] ———, *Linear stability analysis of rotating spiral waves in excitable media*, Physical review letters **68** (1992), no. 13, 2090–2093.

[Bar94] ———, *Euclidean symmetry and the dynamics of rotating spiral waves*, Physical review letters **72** (1994), no. 1, 164–167.

[BCF+10] M. Boulakia, S. Cazeau, M. Fernández, J-F. Gerbeau, and N. Zemzemi, *Mathematical modeling of electrocardiograms: a numerical study*, Annals of biomedical engineering **38** (2010), no. 3, 1071–1097.

[BDSD12] A. Brandejsky, B. De Saporta, and F. Dufour, *Numerical methods for the exit time of a piecewise-deterministic Markov process*, Advances in Applied Probability **44** (2012), no. 1, 196–225.

[BG06] N. Berglund and B. Gentz, *Noise-induced phenomena in slow-fast dynamical systems: a sample-paths approach*, vol. 246, Springer Berlin, 2006.

[BGT13] M. Boulakia, A. Genadot, and M. Thieullen, *Simulations of Stochastic Partial Differential Equations for Excitable Media using Finite Elements*, Submitted (2013).

[BKT90] D. Barkley, M. Kness, and L. Tuckerman, *Spiral-wave dynamics in a simple model of excitable media: The transition from simple to compound rotation*, Physical Review A **42** (1990), no. 4, 2489–2492.

[BLBMZ12] M. Benaïm, S. Le Borgne, F. Malrieu, and P. Zitt, *Quantitative ergodicity for some switched dynamical systems*, Electronic Communications in Probability **17** (2012), 1–14.

[BLMZ12] M. Benaïm, S. LeBorgne, F. Malrieu, and P-A. Zitt, *Qualitative properties of certain piecewise deterministic Markov processes*, arXiv preprint arxiv:1204.4143. (2012).

[BM08] S. Bonaccorsi and E. Mastrogiacoma, *Analysis of the Stochastic Fitzhugh-Nagumo System*, Infinite Dimensional Analysis, Quantum Probability and Related Topics **11** (2008), no. 03, 427–446.

[BN13] P. Bressloff and J. Newby, *Metastability in a stochastic neural network modeled as a velocity jump Markov process*, arXiv preprint arXiv:1304.6960 (2013).

[BR11] E. Buckwar and M. Riedler, *An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution*, Journal of mathematical biology **63** (2011), no. 6, 1051–1093.

[Bre12] P. Bressloff, *Spatiotemporal dynamics of continuum neural fields*, Journal of Physics A: Mathematical and Theoretical **45** (2012), no. 3, 033001.

[CD10] O. Costa and F. Dufour, *Average continuous control of piecewise deterministic Markov processes*, SIAM Journal on Control and Optimization **48** (2010), no. 7, 4262–4291.

[CD11] ———, *Singular perturbation for the discounted continuous control of piecewise deterministic Markov processes*, Applied Mathematics & Optimization **63** (2011), no. 3, 357–384.

[CDMR12] A. Crudu, A. Debussche, A. Muller, and O. Radulescu, *Convergence of stochastic gene networks to hybrid piecewise deterministic processes*, The Annals of Applied Probability **22** (2012), no. 5, 1822–1859.

[CDR09] A. Crudu, A. Debussche, and O. Radulescu, *Hybrid stochastic simplifications for multiscale gene networks*, BMC systems biology **3** (2009), no. 1, 89.

[CF09] S. Cerrai and M. Freidlin, *Averaging principle for a class of stochastic reaction-diffusion equations*, Probability theory and related fields **144** (2009), no. 1-2, 137–177.

[CW96] C. Chow and J. White, *Spontaneous action potentials due to channel fluctuations*, Biophysical Journal **71** (1996), no. 6, 3013–3021.

[CYY07] Y. Cao, H. Yang, and H. Yin, *Finite element methods for semilinear elliptic stochastic partial differential equations*, Numerische Mathematik **106** (2007), 181–198 (English).

[Dav84] M. Davis, *Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models*, Journal of the Royal Statistical Society. Series B (Methodological) (1984), 353–388.

[Dav93] ———, *Markov Models and Optimization*, vol. 49, Chapman and Hall/CRC, 1993.

[Deb11] A. Debussche, *Weak approximation of stochastic partial differential equations: the nonlinear case*, Mathematics of Computation **80** (2011), no. 273, 89–117.

[DHKR12] M. Doumic, M. Hoffmann, N. Krell, and L. Robert, *Statistical estimation of a growth-fragmentation model observed on a genealogical tree*, arXiv preprint arXiv:1210.3240 (2012).

[DIP+10] S. Doi, J. Inoue, Z. Pan, K. Tsumoto, and M. Tanaka, *Computational Electrophysiology: a First Course in in Silico Medicine*, Springer, 2010.

[DP04] G. Da Prato, *Kolmogorov Equations for Stochastic PDEs*, Birkhäuser Basel, 2004.

[DP09] A. Debussche and J. Printems, *Weak order for the discretization of the stochastic heat equation*, Mathematics of Computation **78** (2009), no. 266, 845–863.

[DPZ92] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge University Press, Cambridge, 1992 (English).

[DRL12] A. Destexhe and M. Rudolph-Lilith, *Neuronal noise*, vol. 8, Springer Science + Business Media, 2012.

[EK86] S. Ethier and T. Kurtz, *Markov Processes: Characterization and Convergence*, John Wiley and Sons, 1986.

[ES70] J. Evans and N. Shenk, *Solutions to axon equations*, Biophysical journal **10** (1970), no. 11, 1090–1101.

[FGC08] A. Faggionato, D. Gabrielli, and M. Crivellari, *Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors*, arXiv preprint arXiv:0808.1910 (2008).

[Fit55] R. FitzHugh, *Mathematical models of threshold phenomena in the nerve membrane*, The bulletin of mathematical biophysics **17** (1955), no. 4, 257–278.

[Fit61] ———, *Impulses and physiological states in theoretical models of nerve membrane*, Biophysical journal **1** (1961), no. 6, 445–466.

[Fit69] R. Fitzhugh, *Mathematical models of excitation and propagation in nerve*, Biological Engineering, 1969.

[FL94] R. Fox and Y. Lu, *Emergent collective behavior in large numbers of globally coupled independently stochastic ion channels*, Physical Review E **49** (1994), no. 4, 3421.

[FL07] A. Faisal and S. Laughlin, *Stochastic simulations on the reliability of action potential propagation in thin axons*, PLoS computational biology **3** (2007), no. 5, e79.

[FM13] O. Faugeras and J. MacLaurin, *A large deviation principle for networks of rate neurons with correlated synaptic weights*, arXiv preprint arXiv:1302.1029 (2013).

[FWL05] A. Faisal, J. White, and S. Laughlin, *Ion-channel noise places limits on the miniaturization of the brain's wiring*, Current Biology **15** (2005), no. 12, 1143–1149.

[GM05] I. Gyöngy and A. Millet, *On discretization schemes for stochastic evolution equations*, Potential Analysis **23** (2005), no. 2, 99–134.

[GMV12] L. Goudenège, D. Martin, and G. Vial, *High Order Finite Element Calculations for the Cahn-Hilliard Equation*, Journal of Scientific Computing **52** (2012), no. 2, 294–321.

[Gor12] D. Goreac, *Viability, invariance and reachability for controlled piecewise deterministic Markov processes associated to gene networks*, ESAIM: Control, Optimisation and Calculus of Variations **18** (2012), no. 2, 401–426.

[GSB11] J. Goldwyn and E. Shea-Brown, *The what and where of adding channel noise to the Hodgkin-Huxley equations*, PLoS computational biology **7** (2011), no. 11, e1002247.

[GT12] A. Genadot and M. Thieullen, *Averaging for a fully coupled piecewise-deterministic Markov process in infinite dimensions*, Advances in Applied Probability **44** (2012), no. 3, 749–773.

[GT13a] ______, *Multiscale Piecewise Deterministic Markov Process in Infinite Dimension: Central limit theorem*, arXiv preprint arXiv:1211.1894 (2013).

[GT13b] ______, *On the quantitative ergodicity of some infinite dimensional switching systems and application to averaging*, Submitted (2013).

[Hai08] M. Hairer, *Ergodic Theorem for Infinite Dimensional Systems*, Oberwolfach Reports (2008).

[Hen81] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, vol. 840, Springer Berlin, 1981.

[HH52] A. Hodgkin and A. Huxley, *A quantitative description of membrane current and its application to conduction and excitation in nerve*, The Journal of physiology **117** (1952), no. 4, 500.

[Hil84] B. Hille, *Ionic channels of excitable membranes*, vol. 174, Sinauer Associates Sunderland, MA, 1984.

[Hin02] R. Hinch, *An analytical study of the physiology and pathology of the propagation of cardiac action potentials*, Progress in Biophysics and Molecular Biology **78** (2002), no. 1, 45 – 81.

[HLHOP] F. Hecht, A. Le Hyaric, K. Ohtsuka, and O. Pironneau, *Freefem++, finite elements software.*

[HRW12] M. Hairer, M. Ryser, and H. Weber, *Triviality of the 2D stochastic Allen-Cahn equation*, Electronic Journal of Probability **17** (2012), 1–14.

[Jac05] M. Jacobsen, *Point Process Theory and Applications: Marked Point and Piecewise Deterministic Processes*, Birkhäuser Boston, 2005.

[JC06] P. Jordan and D. Christini, *Cardiac Arrhythmia*, John Wiley & Sons, Inc., 2006.

[Jen09] A. Jentzen, *Pathwise numerical approximations of SPDEs with additive noise under non-global Lipschitz coefficients*, Potential Analysis **31** (2009), no. 4, 375–404.

[JR10] A. Jentzen and M. Röckner, *A milstein scheme for SPDEs*, arXiv preprint arXiv:1001.2751 (2010).

[JS87] J. Jacod and A. Shiryaev, *Limit theorems for stochastic processes*, vol. 288, Springer-Verlag Berlin, 1987.

[Kee80] J. Keener, *Waves in excitable media*, J. App. Math. **39** (1980), no. 3, 528–548.

[KLL10] M. Kovács, S. Larsson, and F. Lindgren, *Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise*, Numerical Algorithms **53** (2010), no. 2-3, 309–320.

[KLNS11] P. Kloeden, G. Lord, A. Neuenkirch, and T. Shardlow, *The exponential integrator scheme for stochastic partial differential equations: Pathwise error bounds*, Journal of Computational and Applied Mathematics **235** (2011), no. 5, 1245–1260.

[KR13] C. Kuehn and M. Riedler, *Large Deviations for Nonlocal Stochastic Neural Fields*, arXiv preprint arXiv:1302.5616 (2013).

[Lam86] L. Lamberti, *Solutions to the Hodgkin-Huxley equations*, Applied Mathematics and Computation **18** (1986), no. 1, 43–70.

[LGONSG04] B. Lindner, J. Garcia-Ojalvo, A. Neiman, and L. Schimansky-Geier, *Effects of noise in excitable systems*, Physics Reports **392** (2004), no. 6, 321 – 424.

[LP13] A. Löpker and Z. Palmowski, *On time reversal of piecewise deterministic Markov processes*, Electron. J. Probab **18** (2013), no. 13, 1–29.

[LS13] E. Luçon and W. Stannat, *Mean field limit for disordered diffusions with singular interactions*, arXiv preprint arXiv:1301.6521 (2013).

[LT10] G. Lord and A. Tambue, *A modified semi-implict Euler-Maruyama scheme for finite element discretization of SPDEs*, arXiv preprint arXiv:1004.1998 (2010).

[LT12] ______, *Stochastic exponential integrators for the finite element discretization of SPDEs for multiplicative and additive noise*, IMA Journal of Numerical Analysis (2012).

[Mé84] M. Métivier, *Convergence faible et principe d'invariance pour des martingales à valeurs dans des espaces de Sobolev*, Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques **20** (1984), no. 4, 329–348 (fre).

[ML81] C. Morris and H. Lecar, *Voltage oscillations in the barnacle giant muscle fiber*, Biophysical journal **35** (1981), no. 1, 193–213.

[MS03] C. Mitchell and G. Schaeffer, *A two-current model for the dynamics of cardiac membrane*, Bulletin of Mathematical Biology **65** (2003), no. 5, 767 – 793.

[NAY62] J. Nagumo, S. Arimoto, and S. Yoshizawa, *An active pulse transmission line simulating nerve axon*, Proceedings of the IRE **50** (1962), no. 10, 2061–2070.

[PS08] G. Pavliotis and A. Stuart, *Multiscale methods: averaging and homogenization*, Springer, 2008.

[PTW10] K. Pakdaman, M. Thieullen, and G. Wainrib, *Fluid limit theorems for stochastic hybrid systems with application to neuron models*, Advances in Applied Probability **42** (2010), no. 3, 761–794.

[PTW12] ———, *Asymptotic expansion and central limit theorem for multiscale piecewise-deterministic Markov processes*, Stochastic Processes and their Applications (2012).

[PZ07] S. Peszat and J. Zabczyk, *Stochastic Partial Differential Equations with Lévy noise*, Cambridge University Press, 2007.

[RB13] M. Riedler and E. Buckwar, *Laws of large numbers and Langevin approximations for stochastic neural field equations*, The Journal of Mathematical Neuroscience (JMN) **3** (2013), no. 1, 1–54.

[Rie12a] M. Riedler, *Almost sure convergence of numerical approximations for Piecewise Deterministic Markov Processes*, Journal of Computational and Applied Mathematics (2012).

[Rie12b] ———, *Spatio-temporal Stochastic Hybrid Models of Biological Excitable Membranes*, Ph.D. thesis, 2012.

[Rin90] J. Rinzel, *Discussion: Electrical excitability of cells, theory and experiment: Review of the Hodgkin-Huxley foundation and an update*, Bulletin of Mathematical Biology **52** (1990), no. 1-2, 3–23 (English).

[RK73] J. Rinzel and J. Keller, *Traveling wave solutions of a nerve conduction equation*, Biophysical Journal **13** (1973), no. 12, 1313–1337.

[RR04] M. Renardy and R. Rogers, *An introduction to partial differential equations*, vol. 13, Springer Verlag, 2004.

[RT83] P. Raviart and J. Thomas, *Introduction à l'analyse numérique des équations aux dérivées partielles*, Masson, 1983.

[RT13] M. Riedler and M. Thieullen, *Spatio-Temporal Hybrid (PDMP) Models: Central Limit Theorem and Langevin Approximation for Global Fluctuations. Application to Electrophysiology*, arXiv preprint arXiv:1304.5651 (2013).

[RTW12] M. Riedler, M. Thieullen, and G. Wainrib, *Limit theorems for infinite-dimensional piecewise deterministic Markov processes. Applications to stochastic excitable membrane models*, Electronic Journal of Probability **17** (2012), 1–48.

[RW07] J. Rubin and M. Wechselberger, *Giant squid-hidden canard: the 3D geometry of the Hodgkin-Huxley model*, Biological Cybernetics **97** (2007), no. 1, 5–32.

[Sac04] F. Sachse, *Computational cardiology: modeling of anatomy, electrophysiology, and mechanics*, vol. 2966, Springer Verlag, 2004.

[San02] B. Sandstede, *Stability of travelling waves, Handbook of dynamical systems, vol. 2, 983–1055*, 2002.

[SBB+12] L. Squire, D. Berg, F. Bloom, S. DuLac, A. Ghosh, and N. Spitzer, *Fundamental neuroscience*, Academic Press, 2012.

[Sha05] T. Shardlow, *Numerical simulation of stochastic PDEs for excitable media*, Journal of computational and applied mathematics **175** (2005), no. 2, 429–446.

[TJ10] H. Tuckwell and J. Jost, *Weak noise in neurons may powerfully inhibit the generation of repetitive spiking but not its propagation*, PLoS computational biology **6** (2010), no. 5, e1000794.

[TK09] M. Tyran-Kamińska, *Substochastic semigroups and densities of piecewise deterministic Markov processes*, Journal of Mathematical Analysis and Applications **357** (2009), no. 2, 385–402.

[Wag80] D. Wagner, *Survey of measurable selection theorems: An update*, Measure Theory Oberwolfach 1979 (Dietrich Kölzow, ed.), Lecture Notes in Mathematics, vol. 794, Springer Berlin Heidelberg, 1980, pp. 176–219.

[Wai10] G. Wainrib, *Randomness in neurons : a multiscale probabilistic analysis*, Ph.D. thesis, 2010.

[Wal81] J. Walsh, *A stochastic model of neural response*, Advances in applied probability (1981), 231–281.

[Wal05] ———, *Finite element methods for parabolic stochastic PDE's*, Potential Analysis **23** (2005), no. 1, 1–43.

[WKAK98] J. White, R. Klink, A. Alonso, and A. Kay, *Noise from voltage-gated ion channels may influence neuronal dynamics in the entorhinal cortex*, Journal of neurophysiology **80** (1998), no. 1, 262–269.

[WR12] W. Wang and A. Roberts, *Average and deviation for slow-fast stochastic partial differential equations*, Journal of Differential Equations (2012).

[WRK00] J. White, J. Rubinstein, and A. Kay, *Channel noise in neurons*, Trends in neurosciences **23** (2000), no. 3, 131–137.

[WTP12] G. Wainrib, M. Thieullen, and K. Pakdaman, *Reduction of stochastic conductance-based neuron models with time-scales separation*, Journal of computational neuroscience **32** (2012), no. 2, 327–346.

[Yan05] Y. Yan, *Galerkin finite element methods for stochastic parabolic partial differential equations*, SIAM journal on numerical analysis **43** (2005), no. 4, 1363–1384.

[YZ98] G. Yin and Q. Zhang, *Continuous-time Markov chains and applications*, Springer, 1998.

[YZ09] G. Yin and C. Zhu, *Hybrid switching diffusions: properties and applications*, vol. 63, Springer, 2009.

[Zam00] L. Zambotti, *An analytic approach to existence and uniqueness for martingale problems in infinite dimensions*, Probability Theory and Related Fields **118** (2000), no. 2, 147–168.